



***SISTEMAS DE VIGILÂNCIA EPIDEMIOLÓGICA
E INFORMAÇÃO GEOGRÁFICA***

***O sistema de apoio a emergências em saúde pública da
Direcção-Geral da Saúde***

Miguel André Fouto Pinho de Oliveira

Dissertação apresentada como requisito parcial para obtenção do
grau de Mestre em Ciência e Sistemas de Informação Geográfica

Instituto Superior de Estatística e Gestão de Informação
da Universidade Nova de Lisboa

**SISTEMAS DE VIGILÂNCIA EPIDEMIOLÓGICA
E INFORMAÇÃO GEOGRÁFICA**

**O sistema de apoio a emergências em saúde pública da
Direcção-Geral da Saúde**

Dissertação orientada por:
Professor Doutor Pedro Cabral

Agosto de 2008

AGRADECIMENTOS

Gostaria de demonstrar o meu agradecimento a todas as pessoas que acompanharam a evolução deste trabalho e que de alguma forma me apoiaram na sua preparação.

Um agradecimento especial vai para o Professor Doutor Pedro Cabral, pela sua orientação, disponibilidade e partilha de conhecimentos, que muito contribuíram para o aperfeiçoamento deste trabalho.

Agradeço ao Director-Geral de Saúde, Dr. Francisco George, por proporcionar as condições para a realização deste trabalho na Direcção-Geral da Saúde e também ao Dr. Mário Carreira, às minhas colegas de projecto Carla Cardoso e Tatiana Plantier, e a todas as colegas do Núcleo de Apoio à Informática, pela constante simpatia e apoio.

Deixo igualmente uma palavra de gratidão incontornável aos meus amigos António Nunes, Hugo Carrão e Vasco Nunes, pela constante ajuda e conselhos que sempre se disponibilizaram a prestar. Agradeço também ao meu colega de mestrado e amigo Eric Noronha Vaz, pelo apoio proporcionado e pelas múltiplas tertúlias, que se revelaram uma boa forma de descansar o espírito. Deixo igualmente uma palavra de agradecimento ao Eng.º Roberto Henriques, pela amabilidade com que me disponibilizou e introduziu ao funcionamento da aplicação *soMGis*.

Acima de tudo, agradeço à Cristina pelo seu constante carinho, apoio e crítica construtiva ao longo de todo este trilha, e aos meus Pais, sem os quais esta caminhada não teria sido possível.

A todos, um muito bem-haja.

*"A verdadeira viagem de descoberta consiste, não em contemplar novas paisagens,
mas em possuir novos olhos"*

Marcel Proust

SISTEMAS DE VIGILÂNCIA EPIDEMIOLÓGICA E INFORMAÇÃO GEOGRÁFICA

O sistema de apoio a emergências em saúde pública da

Direcção-Geral da Saúde

RESUMO

Esta tese pretende ilustrar o estado da arte, o potencial e a aplicação das ferramentas de informação geográfica na gestão da Saúde Pública. A este nível, existem já muitas áreas de actividade, e uma delas consiste na integração de Sistemas de Informação Geográfica (SIG) em sistemas dedicados de vigilância epidemiológica, no sentido de auxiliar os responsáveis de Saúde Pública na tomada de decisões. Vários países possuem actualmente sistemas deste tipo em operação, actuando a diferentes escalas espaciais e temporais, e com diferentes graus de prioridade.

Para além do enquadramento teórico a estes temas, apresenta-se o desenvolvimento, na Direcção-Geral da Saúde (DGS), de uma infra-estrutura de gestão de dados espaciais em Saúde Pública, orientada essencialmente para a vigilância epidemiológica, denominada *Sistema de Suporte a Emergências em Saúde Pública (SSESP)*. Esta consiste num ambiente multiplataforma que conjuga tecnologia de bases de dados relacionais, SIG e mecanismos de disponibilização de conteúdos na Internet. Esta arquitectura sustenta o fornecimento de resultados actualizados com periodicidade diária e semanal a entidades autorizadas (e.g. delegados de saúde), através da Internet. São também referidos estudos pontuais, ou realizados internamente à DGS, que tiram partido dos dados armazenados no sistema e respectivas funcionalidades.

Os resultados da implementação do SSESP são encorajadores, estando já a maior parte da infra-estrutura e funcionalidades planeadas operacional. Subsistem ainda limitações a ultrapassar, sendo também discutidas perspectivas para a sua evolução.

EPIDEMIOLOGICAL SURVEILLANCE SYSTEMS AND GEOGRAPHIC INFORMATION

The public health emergencies support system of the Portuguese General Directorate for Health

ABSTRACT

This thesis aims to display the state of the art, the potential and the application of geographic information tools in Public Health management. This subject already includes many areas of study, one of which deals with the integration of Geographic Information Systems (GIS) in epidemiological surveillance systems, with the objective of aiding Public Health officials in decision-making. Some of these systems are already operational in several countries, acting in various spatial and temporal scales, and with different levels of priority.

Besides providing theoretical support to these themes, this thesis also describes the development of an infrastructure for the management of Public Health spatial data in the Portuguese General Directorate of Health, baptized *Public Health Emergencies Support System (SSESP)*, and essentially aimed at performing epidemiological surveillance tasks. This is a multiplatform environment that brings together relational database technology, GIS and software that supports the availability of contents over the Internet. This architecture makes it possible to supply daily and weekly updated results to health officials through the Internet. Complementary studies that use the accumulated data and tools made available by the system are also referred, although not performed on a regular basis.

Satisfactory results were obtained with the implementation of SSESP. Most of the planned infrastructure and functionalities are already operational. Some of the system's present handicaps and evolutionary perspectives are also discussed.

PALAVRAS-CHAVE

Integração de Sistemas

Internet

Sistemas de Informação Geográfica

Vigilância Epidemiológica

KEYWORDS

Systems Integration

Internet

Geographic Information Systems

Epidemiological Surveillance

ACRÓNIMOS

- BGRI** - Base Geográfica de Referenciação de Informação
- CAOP** - Carta Administrativa Oficial de Portugal
- CDC** - Centers for Disease Control
- CDG** - Conjuntos de Dados Geográficos
- CIG** - Ciência da Informação Geográfica
- CNPHI** - Canadian Network for Public Health Intelligence
- DGS** - Direcção-Geral da Saúde
- ECDC** - European Centre for Disease Prevention and Control
- EMPRES** - Emergency Prevention System for Transboundary Animal and Plant Pests and Diseases
- EMPRES-i** - Global Animal Health Information System
- EPA** - Environment Protection Agency
- ER** - Entidade – Relacionamento
- ESRI** - Environmental System Research Institute
- FAO** - Food and Agriculture Organization
- GAM** - Geographic Analysis Machine
- GASA** - Grupo de Análise de Sistemas Ambientais
- GPS** - Global Positioning System
- GSDI** - Global Spatial Data Infrastructure
- HPAI** - Highly Pathogenic Avian Influenza
- ICD** - International Classification of Diseases
- IGP** - Instituto Geográfico Português
- IIS** – Internet Information Services
- INE** - Instituto Nacional de Estatística
- INSPIRE** - INfrastructure for SPatial InfoRmation in Europe
- INSPQ** - Institut National de Santé Publique du Quebec
- ISPHM-WNV** - Integrated System for Public Health Monitoring of the West Nile Virus

ODBC - Open Database Connectivity

OMS - Organização Mundial de Saúde

PALOP - Países Africanos de Expressão Oficial Portuguesa

PL/SQL - Oracle Procedural Language for SQL

RNA - Rede Neuronal Artificial

RODS - Real-Time Outbreak and Disease Surveillance System

SARS - Severe Acute Respiratory Syndrome

SDI - Spatial Data Infrastructure

SGBD - Sistema Gestor de Bases de Dados

SIDA - Síndrome da Imunodeficiência Adquirida

SIU - Sistema de Informação das Urgências Hospitalares e Centros de Saúde

SOM - Self-Organizing Maps

SQL - Structured Query Language

SIG - Sistemas de Informação Geográfica

SINUS - Sistema Informático de Unidades de Saúde

SNIG - Sistema Nacional de Informação Geográfica

SNS - Serviço Nacional de Saúde

SONHO - Sistema de Informação para Gestão de Doentes Hospitalares

SSESP - Sistema de Suporte a Emergências em Saúde Pública

US NSDI - US National Information Infrastructure

WGS84 - World Geodetic System 1984

WNV - West Nile Virus

ÍNDICE DO TEXTO

AGRADECIMENTOS	iii
RESUMO	v
ABSTRACT	vi
PALAVRAS-CHAVE	vii
KEYWORDS	vii
ACRÓNIMOS	viii
ÍNDICE DE TABELAS	xiii
ÍNDICE DE FIGURAS	xiv
1 INTRODUÇÃO	1
1.1 Saúde Pública, Informação Geográfica, Vigilância Epidemiológica no século XXI.....	1
1.2. Objectivos	2
1.3 Metodologia	3
1.4 Organização da dissertação	4
2 SAÚDE PÚBLICA E CIÊNCIA DA INFORMAÇÃO GEOGRÁFICA	6
2.1 Introdução	6
2.2 Ciência da Informação Geográfica e Geografia da Saúde	6
2.3 Evolução histórica das intervenções da CIG na saúde	7
2.4 Bases de dados para a saúde	10
2.5 Principais domínios e áreas de intervenção	12
2.5.1 Geografia da Doença	12
2.5.2 Geografia dos Sistemas de Cuidados de Saúde	21
2.6 Conclusões	22
3 SISTEMAS DE VIGILÂNCIA EPIDEMIOLÓGICA E INFORMAÇÃO GEOGRÁFICA	25
3.1 Introdução	25
3.2 Sistemas de Vigilância Epidemiológica e Informação Geográfica	25
3.3 Domínios e escalas de intervenção	27

3.4 Exemplos de aplicação	28
3.4.1 Vigilância global da Gripe Aviária Altamente Patogénica	29
3.4.2 Vigilância do Vírus do Nilo Ocidental na província do Quebec, Canadá	32
3.5 Conclusões	37
4 MODELAÇÃO DO SISTEMA DE SUPORTE A EMERGÊNCIAS EM SAÚDE PÚBLICA	39
4.1. Introdução	39
4.2 Estabelecimento do estado da arte e análise de benefícios	39
4.3 Definição do problema e objectivos	41
4.4 Especificações técnicas para a implementação do sistema	44
4.4.1 Modelação da base de dados relacional	44
4.4.2 Aquisição da base de dados geográfica	47
4.4.3 Arquitectura geral do sistema	48
4.4.4 Conclusões	50
5 IMPLEMENTAÇÃO DO SISTEMA DE SUPORTE A EMERGÊNCIAS EM SAÚDE PÚBLICA	52
5.1 Introdução	52
5.2 Implementação do repositório de dados alfanuméricos em Oracle 10g	52
5.3 Automatização do processamento de dados	55
5.4 Interfaces de consulta à base de dados	58
5.5 Outras aplicações do SESP	63
5.5.1 Actividades de reforço do apoio à decisão em vigilância epidemiológica	63
5.5.2 Amostragem espacial na avaliação do nível de saúde e acessibilidade aos cuidados de saúde nas comunidades imigrantes	65
5.5.3 Ensaio de realocização de serviços de urgência com recurso a redes neuronais artificiais não supervisionadas	70
5.6 Conclusões e desenvolvimentos futuros do SESP	82
6 CONCLUSÕES	84
6.1 Contexto de desenvolvimento da aplicação	84
6.2 Desenvolvimento e utilidade do SESP	84

6.3 Limitações e melhoramentos para o SESP	85
6.4 Próximos horizontes	87
REFERÊNCIAS BIBLIOGRÁFICAS	91
ANEXOS	101
<i>Anexo 1 - Especificações técnicas das tabelas da base de dados do SESP</i>	<i>102</i>
<i>Anexo 2 – Excerto da sequência de execução das tarefas, Job 1</i>	<i>105</i>

ÍNDICE DE TABELAS

Tabela 1 – Tabelas da base de dados.....	46
Tabela 2 – Metadados da informação geográfica.....	48
Tabela 3 – Resultados dos cenários de realocização ensaiados (habitantes servidos por secção estatística).	80

ÍNDICE DE FIGURAS

Figura 1 – Organização da dissertação.	4
Figura 2 – Bases cartográficas no contexto norte-americano	10
Figura 3 – Modelação do impacto de um agente ambiental na população humana	13
Figura 4 – Análise de <i>clusters</i> com recurso a GAM de Stan Openshaw	14
Figura 5 – Processos de difusão espacial	16
Figura 6 – Picos temporais e evolução das infecções por <i>rotavirus</i> nos EUA	17
Figura 7 – Aspecto da aplicação <i>Body Viewer</i>	18
Figura 8 – Surtos de gripe aviária altamente patogénica entre Janeiro e Junho de 2006	31
Figura 9 – Densidade de surtos de Gripe Aviária Altamente Patogénica.....	32
Figura 10 – Análise das recolhas de corvos mortos na área de Montreal em 2005	35
Figura 11 – Aspecto do <i>EpiScanGIS</i> (área pública)	36
Figura 12 – Modelo Conceptual do SESP.	43
Figura 13 – Modelo Entidade – Relacionamento.	45
Figura 14 – Esquema relacional da base de dados do SESP.	47
Figura 15 – Arquitectura do SESP.....	49
Figura 16 – Geração de tabela em <i>Oracle SQL Developer</i>	53
Figura 17 – Indexação da tabela EURO_URGENCIAS_H.	55
Figura 18 – Programação de <i>Chain</i> em <i>Oracle Enterprise Manager 10g Database Control</i> .56	
Figura 19 – Aspecto de modelação de um <i>Job</i> em ambiente <i>Mind Manager</i>	57
Figura 20 – Página de entrada no sítio na Internet de acesso ao SESP	58
Figura 21 – Esquema geral dos produtos disponibilizados no âmbito do SESP.	59
Figura 22 – Página de visualização dos produtos do SESP	59
Figura 23 – Procura distrital de urgências por 100.000 habitantes em 12-05-2008.....	60
Figura 24 – Evolução da procura, Distrito de Beja, 2007 e 2008	62
Figura 25 – Corredor epidémico nacional semanal, ano de 2008	62
Figura 26 – Cordão sanitário gerado a partir da localização de um foco epidémico.....	64

Figura 27 – Evolução da procura de serviços de urgência, dias 11-1-2007, 05-02-2007 e 01-04-2007, concelhos do distrito de Viana do Castelo.	64
Figura 28 – Concelhos amostrados	65
Figura 29 – Fluxograma das operações realizadas	68
Figura 30 – Validação de um <i>cluster</i> da amostra	69
Figura 31 – Ilustração dos princípios do SOM.....	73
Figura 32 – Área de estudo.....	75
Figura 33 – Opções iniciais da aplicação.....	77
Figura 34 – Final do treino da rede neuronal e parametrização.	78
Figura 35 – Localizações actuais de serviços de urgência e polígonos de <i>thiessen</i>	79
Figura 36 – Localizações alternativas para serviços de urgência e polígonos de <i>thiessen</i> . .	79
Figura 37 – Georreferenciação de um foco epidémico detectado num parque de campismo. .	82
Figura 38 – Níveis de alerta distritais, 05-02-2007.	83

1 INTRODUÇÃO

1.1 Saúde Pública, Informação Geográfica, Vigilância Epidemiológica no século XXI

A crescente consciencialização da importância da detecção e interpretação dos padrões de distribuição espacial dos fenómenos relacionados com a Saúde Pública surge associada a um leque de novas possibilidades tecnológicas, postas à disposição dos profissionais e investigadores desta área.

A dinâmica associada a estas novas tecnologias não apenas inspirou o aparecimento de novos rumos metodológicos, mas veio também tornar possível a aplicação de certas metodologias de análise que se encontravam “no frio”, por ausência de ferramentas capazes de as concretizar, não só no que respeita à sua mera execução, mas em grande parte à capacidade de realizar as análises desejadas em tempo útil.

Entre estas tecnologias surgem, em lugar de destaque, os Sistemas de Informação Geográfica (SIG), cuja utilização trouxe à Geografia não apenas a capacidade de implementar metodologias de análise espacial, que até aí se revelavam pouco exequíveis, mas também novos horizontes enquanto disciplina.

A utilização destas metodologias e ferramentas em Saúde Pública constitui um campo recente de aplicação, existindo já, todavia, um registo apreciável de publicações relacionadas com o uso de informação geográfica nesta área. As contribuições fazem-se já sentir em múltiplas facetas de análise em Saúde Pública, tais como na compreensão da influência que os factores ambientais possuem na Etiologia¹ de certas doenças; na tentativa de melhorar a eficácia da prestação de cuidados de saúde às populações, intervindo na preparação de modelos de localização óptima para infra-estruturas e permitindo direccionar mais eficazmente os meios de intervenção rápida para locais de acidente; no desenvolvimento de sistemas de vigilância epidemiológica, os quais são neste momento alvo de especial atenção, pois numa era dominada pelo poder dos *media*, nunca a amplitude e a frequência das ameaças ao estado de saúde das populações foi tão largamente divulgada (e por vezes exacerbada). Estes sistemas assumem um papel de crescente importância, na medida em que constituem cada vez mais a primeira linha de defesa da Saúde Pública, na eventualidade da ocorrência de uma

¹ Estudo das causas das doenças. Silman e Mcfarlane (2004).

pandemia². No sentido de ampliar as capacidades de monitorização e alerta, as ferramentas destes sistemas devem incluir necessariamente tecnologias de informação geográfica.

Na eminência de uma eventual pandemia de gripe, a qual poderá ser originada por uma mutação do vírus da gripe aviária (H5N1), tornando-o transmissível entre indivíduos da espécie humana, tem-se verificado o desenvolvimento de estratégias e infra-estruturas capazes de dar resposta a esta situação, entre as quais surgem os sistemas de vigilância epidemiológica dedicados a esta doença. Se por um lado, a globalização e as novas tecnologias agem como opositores na contenção de vírus e doenças, por facilitarem a rápida deslocação de pessoas contaminadas e vectores³, por outro lado operam, irrevogavelmente, como aliados, permitindo a comunicação e transmissão de conhecimento em tempo real. É assim, mais importante do que nunca, aproveitar também as tecnologias emergentes para assegurar a colaboração e o funcionamento “em rede” dos diversos actores intervenientes em Saúde Pública, garantindo o intercâmbio de ideias e conhecimento para que seja possível responder eficazmente a qualquer tipo de alerta.

1.2. Objectivos

Esta dissertação compreende seis objectivos principais, sendo um deles de carácter primordial, e do qual decorrem os restantes. O objectivo essencial desta dissertação prende-se com salientar e demonstrar as mais-valias do recurso a informação geográfica e ferramentas relacionadas, no apoio à decisão em Saúde Pública em geral e, mais especificamente, na vigilância epidemiológica levada a cabo pela Direcção-Geral da Saúde (DGS). Este objectivo essencial subdivide-se assim em vários outros:

- Introduzir e tipificar as áreas de utilização dos sistemas de informação geográfica em Saúde Pública e dar exemplos de casos de estudo;

² Epidemia que ocorre globalmente, ou numa área muito vasta, atravessando fronteiras internacionais e afectando geralmente um grande número de pessoas. EPA (2008).

³ Insecto ou outro animal que transporta uma doença, podendo transmiti-la a seres humanos. Colin (2005).

- Estabelecer o estado da arte da aplicação de tecnologias de informação geográfica em Saúde Pública e uma breve evolução dos casos de estudo efectuados até ao presente;
- Apresentar e definir a actividade de vigilância epidemiológica, indicando também sistemas relacionados;
- Referir o planeamento e montagem, na DGS, de um sistema de vigilância epidemiológica, denominado *Sistema de Suporte a Emergências em Saúde Pública (SSESP)*, que integra componentes muito diversos, entre os quais um sistema de informação geográfica.
- Indicar as actividades levadas a cabo pelo referido sistema, as quais se dividem entre as operações diárias de monitorização e eventual emissão de alertas epidemiológicos, e o seu funcionamento como repositório de dados, fornecendo apoio à realização de outras análises e estudos, tanto de carácter retrospectivo como prospectivo.

1.3 Metodologia

A metodologia adoptada compreendeu os seguintes passos:

- Pesquisas bibliográficas e consultas aos vários intervenientes (médicos, técnicos de Saúde Pública, colaboradores, fornecedores de *software* e *hardware*, entre outros) no sentido de adquirir os conhecimentos que permitiram identificar as funcionalidades a incluir no sistema;
- Planeamento e delineação do modelo a seguir na implementação do sistema em termos dos dados que o alimentam, da sua arquitectura e componentes;
- Diagnóstico das ferramentas de hardware e software necessárias, bem como da infra-estrutura de comunicações que suporta as interacções entre os vários componentes, garantindo a sua integração;
- Implementação do sistema;
- Realização de estudos complementares, relacionados, ou não, com a actividade de vigilância epidemiológica diária, que tiram partido dos dados continuamente armazenados no sistema e das ferramentas que disponibiliza;

- Avaliação geral dos desenvolvimentos produzidos e recomendações futuras.

1.4 Organização da dissertação

Esta dissertação é constituída por seis capítulos (figura 1), sendo este o capítulo introdutório, onde se refere o contexto que motiva o desenvolvimento do presente estudo, os principais objectivos que pretende atingir e a sua organização geral.

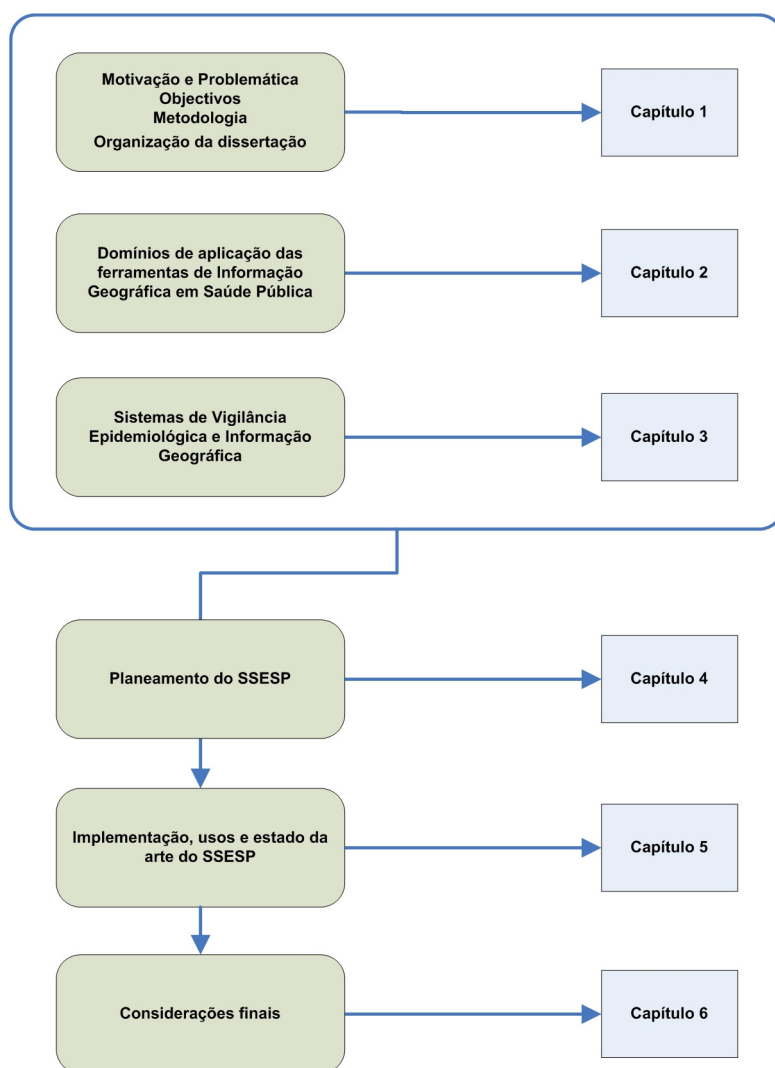


Figura 1 – Organização da dissertação.

No segundo capítulo é traçada uma tipologia e análise do estado da arte dos domínios de aplicação das ferramentas de informação geográfica em Saúde Pública, incluindo a apresentação de exemplos de implementação. O terceiro capítulo constitui uma introdução aos sistemas de vigilância epidemiológica e ao

papel que nestes desempenham as tecnologias de informação geográfica, sendo também apresentados exemplos de alguns sistemas deste tipo actualmente em funcionamento.

Os capítulos iniciais constituem um corpo de estudo que fornece bases e enquadramento teórico para as tarefas desenroladas nos capítulos seguintes.

O quarto capítulo refere-se à fase inicial de planeamento do SESP, incluindo a modelação conceptual do sistema e a definição das especificações técnicas gerais para a sua implementação.

Com o quinto capítulo inicia-se a elaboração da parte prática deste projecto, sendo descrito o desenvolvimento e montagem física do SESP na sua vertente de vigilância epidemiológica diária, e o seu estado da arte. Inclui também a realização de estudos complementares que demonstram as potencialidades e a polivalência desta infra-estrutura, culminando com a apresentação das funcionalidades previstas, mas ainda não implementadas.

O sexto e último capítulo constitui um epílogo do trabalho executado, incluindo vantagens e limitações do sistema implementado e indicando recomendações para o aperfeiçoamento e caminhos futuros a percorrer.

2 SAÚDE PÚBLICA E CIÊNCIA DA INFORMAÇÃO GEOGRÁFICA

2.1 Introdução

Como refere um provérbio chinês: “a mais longa caminhada começa com um passo”. De igual forma, este capítulo fornece o primeiro passo para o entendimento da relação existente entre a prática de Saúde Pública e a Informação Geográfica, considerada aqui como fazendo parte de um conceito mais amplo e recente, a Ciência da Informação Geográfica (CIG). Para tal, a segunda secção expõe o raciocínio que preside à relação entre estas duas áreas, dando origem à *Geografia da Saúde*, a terceira apresenta brevemente a evolução histórica dos estudos e a quarta refere as bases de dados disponíveis e normalmente utilizadas. A quinta secção tipifica as áreas de aplicação e apresenta casos de estudo, e a última secção encerra o capítulo, retirando conclusões.

2.2 Ciência da Informação Geográfica e Geografia da Saúde

Importa definir, antes de mais, em que consiste a CIG e quais os seus elementos em foco. Neste sentido, a definição proposta por Painho (2006) vem fornecer essa estrutura, ao referir-se a esta ciência como um campo de pesquisa experimental na qual o *espaço geográfico* é o objecto de estudo, a *informação geográfica* o princípio refutável e o *sistema de informação geográfica* a ferramenta disciplinar. Cada um destes componentes pode ser equacionado no presente contexto.

No que diz respeito ao estudo do *espaço geográfico*, as intervenções da CIG na saúde enquadram-se no âmbito da *Geografia da Saúde* (reconhecida apenas em 1992), sendo anteriormente posicionadas no âmbito da *Geografia Médica*, designação menos abrangente, mas presente em muita da literatura disponível. Santana (2005) refere esta evolução: “Começando por se confundir com o mero estudo das topografias médicas ou sanitárias, a Geografia Médica transformou-se em Geografia da Saúde, para melhor expressar quer a sua riqueza metodológica e de objectivos, quer a abrangência temática da sua intervenção”. No que diz respeito à abrangência temática, este é um ramo disciplinar de “largo espectro”, que nas últimas três décadas tem abordado um vasto leque de questões, tais como os

factores que influenciam a distribuição e difusão das doenças⁴, a localização dos equipamentos, as áreas de influência e atracção dos serviços de saúde, entre outros.

A *informação geográfica* é aqui abordada em duas vertentes: os dados que a permitem produzir são debatidos na secção 3; a informação e respectivos usos são debatidos no contexto dos casos de estudo apresentados na quarta secção.

Restam os *Sistemas de Informação Geográfica*. Segundo a definição do Environmental System Research Institute (ESRI), que pode ser consultada em ESRI (2006): uma colecção integrada de *software* e dados, utilizados para visualizar e gerir informação acerca de localizações geográficas, analisar relações geográficas e modelar processos espaciais. Um SIG providencia uma estrutura para adquirir e organizar dados espaciais e informação relacionada, para que estes possam ser ilustrados e analisados.

Não obstante a sua frequente interacção com outras ferramentas (tecnologias de processamento digital de imagens de Detecção Remota e pacotes estatísticos, entre outras), esta constitui a ferramenta de eleição das aplicações da CIG na saúde, o ponto de encontro entre os dados e as metodologias para obtenção de informação e conhecimento. Os casos de estudo aqui mencionados têm a ver com a aplicação destas ferramentas.

2.3 Evolução histórica das intervenções da CIG na saúde

Segundo Boulos *et al* (2001), a ideia de que a localização geográfica pode ter influência na saúde é muito anterior ao surgimento da CIG ou até da Informática. A noção de que certas doenças tendem a ocorrer em alguns lugares em detrimento de outros remonta à época de Hipócrates (século terceiro antes de Cristo). Todavia, o primeiro caso de estudo que se aproxima das aplicações actuais foi efectuado em Londres, durante a epidemia de cólera de 1854, por John Snow, médico da época, através de um processo de análise espacial (dado os resultados dependerem da disposição espacial das variáveis em análise). Após registar a localização das habitações das pessoas afectadas pela doença num mapa das ruas do bairro onde se declarou o surto, observou que uma grande parte das ocorrências tinha tido

⁴ Entende-se doença como um estado do organismo em que existem alterações anatómicas ou perturbações funcionais que o afastam das condições normais. ALEA (1999).

lugar nas proximidades de uma determinada bomba de água, dando a entender que o surto tinha origem nesta e levando à decisão do seu encerramento, o que travou a epidemia. Todavia, esta não é considerada uma aplicação SIG, na medida em que não faz uso da maioria dos elementos que definem actualmente um sistema desta natureza. A cronologia destas intervenções inicia-se após o aparecimento da Informática. Segundo Albert *et al.* (2000), os laços entre a saúde e os SIG têm o seu início por volta da década de 80 do século XX, mas ainda de forma embrionária, começando apenas a surgir com alguma força na década de 90. Assim, até 1990 existia apenas um pequeno número de publicações que referem a utilização dos SIG em situações relacionadas com a saúde; entre 1991 e 1994 o número de publicações atingia cerca de uma dúzia por ano, continuando este número a aumentar anualmente entre 1995 e 1997. Os autores recorrem a um modelo de difusão para ilustrar a evolução da pesquisa nesta área, sendo que a difusão se refere ao número de aderentes a uma inovação (neste caso uma publicação) ao longo de um período de tempo. Assim, os resultados neste tipo de investigação começam a surgir no fim da década de 1980 até 1990 (fase 1) e a fase de expansão inicial (fase 2) decorre 1991 e 1997. À data desta publicação, os autores previam que esta fase se prolongasse por alguns anos mais até ser atingida a fase de expansão tardia (fase 3), e que daí até à fase de saturação (fase 4) mediasse uma década ou mais (aproximadamente 2015). Completado o processo de difusão, estas tecnologias seriam parte activa e corrente na investigação em saúde. Em face das dificuldades ainda existentes na aceitação destas ferramentas por parte dos profissionais de saúde, apontadas por Scotch (2006), esta previsão parece ser realista.

Um factor de esperança no desenvolvimento destas aplicações reside no facto de, para além de se verificar o aparecimento de estudos em publicações periódicas generalistas, tanto ao nível de Informação Geográfica (e.g. *International Journal of Geographical Information Science*), como de Saúde Pública e Medicina (*Medical Informatics and Decision Making*, *Malária Journal*, *Medicine*, entre outras), existir já pelo menos uma publicação periódica dedicada ao tema, o *International Journal of Health Geographics*. Este periódico é disponibilizado no sítio na Internet desta publicação (IJHG, 1999), sendo os artigos apresentados de acesso livre e gratuito. Possui já uma dinâmica apreciável, apresentando desde estudos de carácter retrospectivo (Boulos, 2005) até aplicações complexas de estatísticas espacio-

temporais para detecção precoce de surtos⁵ (Takahashi *et al*, 2008) e de Geostatística (Goovaerts e Gebreab, 2008), passando por aplicações de Detecção Remota em Saúde Pública (Tran *et al*, 2008). Possui também a vantagem de se enquadrar na plataforma *BioMed Central*, dedicada à partilha de informação de acesso livre relativa a estudos médicos em geral (BioMed Central, 1999), o que lhe confere flexibilidade na comunicação dentro da área médica. Para além desta publicação periódica, existem também já vários volumes não periódicos publicados, que demonstram a riqueza do campo e as suas potencialidades futuras. Variam entre publicações abrangentes em saúde, SIG e tecnologias relacionadas, como os trabalhos de Gatrell e Löytönen (1998), Albert *et al* (2000), Cromley e McLafferty (2002) e Khan (2003), ou focadas em áreas mais específicas, como cartografar e avaliar riscos para a Saúde Pública (Lawson *et al*, 1999), aplicações de estatística espacial em Saúde Pública (Waller e Gotway, 2004), Epidemiologia Espacial (Elliot *et al*, 2000) e a utilização de métodos estatísticos em Epidemiologia Espacial (Lawson, 2006), entre outras.

Outra perspectiva prende-se com a evolução das ferramentas de software SIG e relacionadas. A maior parte dos estudos de Saúde Pública contendo uma componente de análise espacial recorre a ferramentas de software comercial generalista (e.g. ESRI Arc/INFO), tendência relativamente consistente desde o surgimento da área de aplicação, e que beneficia do facto destas ferramentas se apresentarem, em geral, cada vez mais amigáveis para o utilizador. Por outro lado, as ferramentas SIG de código fonte aberto (software vulgarmente designado *Open Source*) começam a ser utilizadas, existindo já desde aplicações que podem ser instaladas num computador isolado, até outras vocacionadas para suportar a disponibilização de conteúdos geográficos através de sítios na Internet, como se pode observar no desenvolvimento das aplicações *EpiVue* (Yi *et al*, 2008), *ISPHM-WNV* (Gosselin *et al*, 2005) e *EpiScanGIS* (Reinhardt *et al*, 2008), estas duas últimas aprofundadas no capítulo 3.

Por último, entre os desenvolvimentos recentes, destaca-se o surgimento de software SIG especialmente desenvolvido para a aplicação em estudos de Saúde Pública, gratuito ou não, de que já são exemplos:

- a ferramenta de Epidemiologia Espacial *Epi Info*, disponibilizada

⁵ *Surto* corresponde ao aparecimento súbito de uma série de casos de uma doença. Colin (2005).

gratuitamente pelo CDC (*Centers For Disease Control*) em sector dedicado do sítio na Internet desta instituição (CDC, 2008);

- a ferramenta de Epidemiologia Espacial de distribuição comercial *EpiAnalyst* (extensão para ESRI ArcViewGIS 3.x), acerca da qual é possível obter informação em sítio na Internet próprio (REGS, 1999);
- a aplicação de mapeamento e vigilância epidemiológica *HealthMapper*, disponibilizada gratuitamente pela Organização Mundial de Saúde (OMS) em sector dedicado do sítio na Internet desta instituição (OMS, 2008);
- o software *SIGEpi*, desenvolvido pela Organização Pan-Americana de Saúde, distribuído sob licença em sector dedicado do sítio na Internet desta instituição (Organización Panamericana de Salud, 2003).

2.4 Bases de dados para a saúde

Existe um grupo relativamente vasto de dados que podem ser integrados num SIG e utilizados em aplicações deste tipo. No que respeita a bases de dados espaciais, Cromley e McLafferty (2002) dividem estes dados em vários tipos, entre os quais os que designam como *foundation databases*, constituindo bases cartográficas que podem ser usadas em várias áreas de aplicação e permitindo, através de técnicas variadas (como por exemplo a *geocodificação*), referenciar geograficamente dados não espaciais. A figura 2 contém alguns exemplos no contexto norte-americano.

Database	Scale	Source	Type
Digital orthophoto quadrangles	1:12,000–1:25,000	U.S. Geological Survey	Raster
Satellite imagery	1:25,000–1:100,000	Commercial	Raster
Digital line graphs	1:24,000 and 1:100,000	U.S. Geological Survey	Vector
TIGER	≥ 1:100,000	U.S. Bureau of the Census	Vector
Cadastral	1:200–1:12,000	Local government	Vector

Figura 2 – Bases cartográficas no contexto norte-americano (Cromley e McLafferty, 2002).

As fontes para obtenção destes dados são bastante diversas. Boulos (2004) discute o aparecimento recente das infra-estruturas de dados espaciais (*Spatial Data Infrastructures - SDI*) em algumas áreas do mundo desenvolvido, entre as quais se contam, nos Estados Unidos da América, a US NSDI (*US National Information*

Infrastructure), na Europa a iniciativa INSPIRE (*IN*frastructure for *SP*atial *Info*Rmation in Europe) e, a um nível global, a GSDI (*Global Spatial Data Infrastructure*). Em Portugal é de referir o SNIG (Sistema Nacional de Informação Geográfica). Estas infra-estruturas são importantes fontes de dados espaciais também para a área da saúde, a que não é alheia a disponibilidade, ainda que por vezes parcial, destes dados na Internet. Uma descrição bastante completa de tipos e fontes de dados, embora vocacionada para as doenças cancerígenas nos EUA, pode ser consultada em Boscoe *et al.* (2004). É ainda de referir a importância e crescente dinâmica da utilização de dados obtidos por Detecção Remota. Albert *et al.* (2000) inclui uma discussão muito completa destes dados e suas aplicações em saúde. A título de exemplo, Ward *et al.* (2004) referem o uso de imagens de satélite e de fotografia aérea em estudos relacionados com o cancro para detectar habitações situadas em áreas agrícolas, com o objectivo de estimar a exposição dos seus habitantes aos pesticidas.

Existe ainda a possibilidade das instituições de Saúde Pública produzirem os seus próprios dados espaciais, embora esta situação seja muito menos frequente, ocorrendo geralmente em pequenos volumes de dados, em estudos que obedecem a uma lógica de projecto, ou ainda quando os dados não se encontram de todo disponíveis ou actualizados para uso em tempo útil.

Outro género fundamental diz respeito a bases de dados alfanuméricos (não espaciais) directamente relacionados com a saúde, vários dos quais podem já ser facilmente integrados em ambiente SIG e associados aos dados espaciais atrás referidos, sendo potencialmente espacializáveis. Entre estes incluem-se as estatísticas de natalidade e mortalidade e os dados relacionados com a morbilidade (número de casos registados de uma determinada doença), destacando-se aqueles recolhidos por sistemas de vigilância epidemiológica, que monitorizam a distribuição e tendências dos dados da morbilidade e mortalidade recolhidos para um universo populacional específico e para uma dada região; um exemplo português é o sistema de vigilância epidemiológica desenvolvido pela DGS, no qual são recolhidos, integrados no mesmo sistema e interpretados diariamente, dados referentes às idas às urgências em hospitais (provenientes do Sistema de Informação para Gestão de Doentes Hospitalares - SONHO) e centros de saúde (provenientes do Sistema Informático de Unidades de Saúde - SINUS). Em termos de fontes, estes dados são em geral recolhidos e disponibilizados por entidades públicas.

2.5 Principais domínios e áreas de intervenção

É conveniente, e cada vez mais frequente, dividir estas intervenções em dois grandes domínios, delineados por Boulos *et al*, (2001): a Geografia da Doença (*Disease Geography*) e a Geografia dos Sistemas de Cuidados de Saúde (*Geography of Health Care Systems*).

2.5.1 Geografia da Doença

Boulos *et al*, (2001) fornece uma definição abrangente para este ramo de aplicações; “A geografia da doença, [...] cobre a exploração, descrição e modelação da incidência⁶ espaço-temporal das doenças e fenómenos ambientais relacionados, a detecção e análise de padrões e *clusters*⁷ das doenças, a análise das causas e a geração de novas hipóteses.” Esta definição engloba várias actividades de pesquisa, as quais são tipificadas de forma diferenciada, dependendo do autor. Tentando conjugar as contribuições de Cromley *et al* (2002) e Albert *et al*, (2000), esta área corresponde à prevenção e controle de problemas de saúde relacionados com o ambiente. Tipicamente, estes problemas envolvem a acção de agentes que produzem efeitos adversos na saúde humana, podendo ser classificados como *físicos* (exemplo da radiação ultravioleta), *químicos* (por exemplo o chumbo) ou *biológicos*. Estes últimos abrangem a análise do risco e propagação de doenças infecciosas (transmissíveis de pessoa para pessoa) e das doenças transmitidas por vectores (*vector-borne diseases*). Geralmente as populações humanas entram em contacto com estes agentes, presentes na atmosfera, alimentos ou bebidas, através da sua inalação, ingestão ou contacto físico, no contexto do ambiente natural ou urbano. Os estudos mais frequentes na literatura referem-se aos factores *físicos* e *químicos*. O processo pelo qual um agente ambiental deste tipo produz um efeito adverso na população humana pode ser modelado como uma sequência do tipo: risco ambiental → exposição → efeito (figura 3).

⁶ Incidência corresponde ao número de ocorrências de uma doença numa determinada população ao longo de um período de tempo específico. Beaglehole *et al*, (1993).

⁷ Concentrações atípicas da ocorrência de problemas de Saúde Pública no tempo e no espaço. Cromley e McLafferty, (2002).

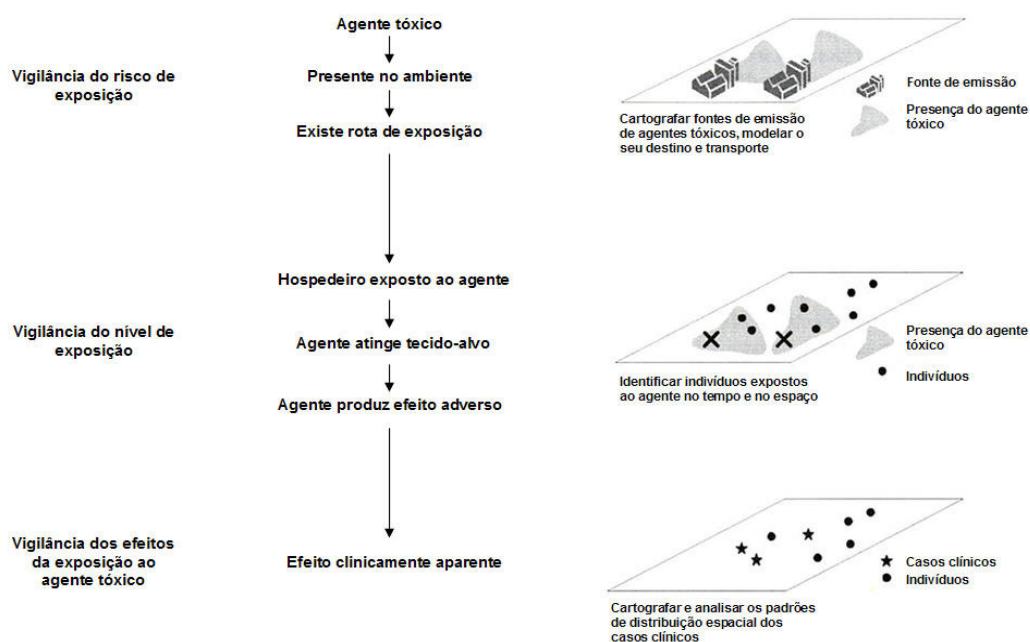


Figura 3 – Modelação do impacto de um agente ambiental na população humana (Cromley e McLafferty, 2002).

A modelação destes processos requer dados referentes a cada um dos passos, isto é, o tipo de risco ambiental (e.g. radiação nuclear), a sua quantidade e origem, indicadores da qualidade ambiental nos locais em que as pessoas vivem e levam a cabo as suas actividades diárias e efeitos adversos relacionados com a exposição aos riscos ambientais. Todos os dados terão de ser obtidos a partir de várias fontes e integrados; nesta integração, as dimensões espacial e temporal são as questões fundamentais para a modelação eficaz da sequência acima descrita. É neste contexto que surgem as aplicações SIG típicas em saúde ambiental, como seja a visualização de fontes de contaminantes ambientais e modelação das zonas de contaminação associadas a estas fontes, a análise das variações da qualidade ambiental através de valores medidos em estações de monitorização e o estudo da distribuição das populações em risco no território e dos padrões espaciais dos efeitos adversos na saúde. Tanto Cromley e McLafferty (2002) como Albert *et al* (2000) incluem uma discussão extensa dos métodos, tipologia e problemas associados a estes estudos. Um exemplo de aplicação, referido em ambas as publicações, tem a ver com a reciclagem de combustível nuclear utilizado numa unidade destinada a este efeito, localizada em *Sellafield*, na região de *West Cumbria* (Reino Unido). A suspeita de que o funcionamento desta unidade possa ter provocado um aumento dos casos de leucemia infantil na região levou Stan

Openshaw e colaboradores a experimentarem, em 1988, uma nova técnica de análise de *clusters* em ambiente SIG, denominada *Geographic Analysis Machine* (GAM), recorrendo a registos de ocorrências de cancro em Manchester e Newcastle (Cromley e McLafferty, 2002). A figura 4 ilustra os resultados da análise.

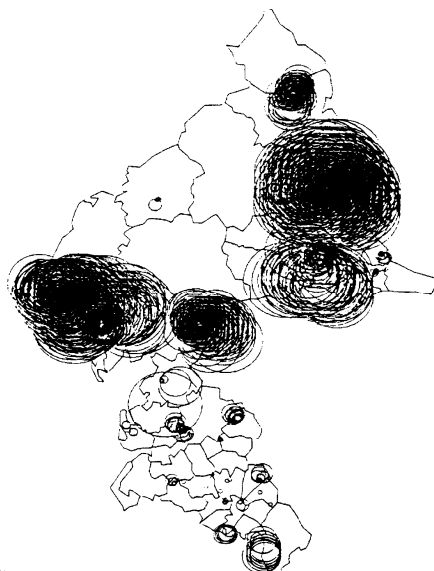


Figura 4 – Análise de *clusters* com recurso a GAM de Stan Openshaw (Cromley e McLafferty, 2002).

Esta técnica tem a particularidade de desenhar aleatoriamente uma multiplicidade de círculos sobre a área de estudo. A fase seguinte consistiu em contar o número de ocorrências de cancro circunscritas por cada círculo e isolar aqueles que continham um número de casos considerado excessivo. A previsão feita para a área de estudo através de métodos estatísticos indicava 173 círculos, mas foram detectados com esta metodologia cerca de 1972 círculos com um número de casos em excesso (as questões relacionadas com a análise de *clusters* voltarão a ser abordadas no final desta secção). No que diz respeito à acção de agentes biológicos, a análise do risco e da propagação de doenças infecciosas tem vindo a assumir uma importância crescente. Doenças infecciosas são todas aquelas causadas por agentes biológicos de dimensão microscópica (microrganismos), designados por *agentes*. Estes propagam-se entre *hospedeiros* humanos ou animais por *transmissão directa* (*doenças contagiosas*) ou através de um terceiro interveniente, designado *vector*. Entre as doenças contagiosas mais sujeitas a análise geográfica, encontram-se algumas que são merecedoras de grande preocupação por parte dos profissionais de Saúde Pública. Algumas das mais

mediáticas são as sexualmente transmissíveis, como a síndrome da imunodeficiência adquirida (SIDA) e a gonorreia. Outras são mais comuns, mas constituem igualmente motivo de preocupação, tais como a tuberculose, gripe e tosse convulsa. A característica mais distinta destas doenças é o facto de se propagarem ao longo do tempo, de pessoa para pessoa e de localização para localização, assumindo muitas vezes dimensões *epidémicas* quando se propagam a uma escala não global mas significativa (nacional ou regional) ou *pandémicas*, quando se propagam globalmente.

O modo de transmissão tem uma importância crítica na utilização dos SIG na avaliação destas doenças, influenciando o tipo de perguntas de carácter geográfico sobre a doença, as análises espaciais efectuadas e os temas geográficos que devem ser incluídos no SIG. Por exemplo, a modelação de doenças transmitidas através do ar tem a ver com os padrões geográficos das interacções humanas, relacionadas com condições de alojamento, lugares onde se juntam multidões e contactos pessoais que ocorrem em instalações, tais como escolas e locais de trabalho. Procurar estabelecer os padrões geográficos de interacção entre pessoas infectadas e possíveis hospedeiros constitui o passo seguinte para compreender quando e como se propagam estas doenças, cujo desenvolvimento ocorre através de processos de *difusão espacial*, a qual diz respeito ao movimento de fenómenos (por exemplo pessoas, ideias e doenças) ao longo do espaço e do tempo. Todas as variedades de difusão se manifestam aqui; por *contágio*, *hierárquica* e *em rede* (figura 5). A *difusão por contágio* ocorre naturalmente, dada a interacção entre pessoas ter maior probabilidade de ocorrer entre as que estão próximas do que as que se encontram distantes. Das limitações à mobilidade das pessoas provocadas pela idade, baixo rendimento ou fraco acesso aos meios de transporte resultam em geral padrões muito localizados de interacção. Mas por vezes as doenças parecem “saltar” entre locais afastados, em vez de se difundirem por contágio. Nos processos de *difusão hierárquica*, as doenças propagam-se ao longo da hierarquia urbana, passando directamente de um núcleo urbano principal para as suas áreas suburbanas ou cidades-satélite. Por último, é provável que percorram caminhos coincidentes com as redes rodoviárias ou ferroviárias, de acordo com as deslocações pendulares ou sazonais de grande número de pessoas; neste caso será uma *difusão em rede*. Em geral, todavia, o que acaba por ocorrer na actualidade é uma mistura complexa destes três tipos, ao invés do que acontecia séculos atrás, quando a maior parte destes fenómenos se difundia por *contágio*.

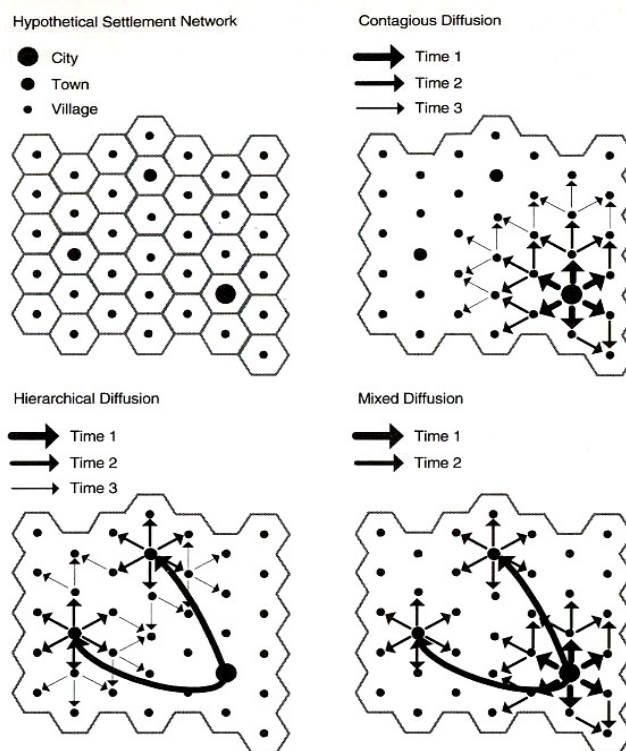


Figura 5 – Processos de difusão espacial (Cromley e McLafferty, 2002).

Uma questão central na monitorização de doenças contagiosas e no planeamento das intervenções necessárias prende-se com a compreensão dos padrões de propagação não apenas no espaço, mas também no *tempo*. Em estudos desta natureza, torna-se por vezes importante, no caso de doenças de origem viral como a gripe, com padrões sazonais definidos, estabelecer quando e como se dão os picos de ocorrências. Cromley e McLafferty (2002) referem que, num estudo realizado nos EUA, os picos sazonais de ocorrências de *rotavirus* (vírus responsável por certos casos de gastroenterite nas crianças) em cada Estado, foram cartografados através da localização pontual de 69 laboratórios em 42 Estados; a cada ponto foi atribuído o número da semana do ano (1 – 52) na qual foi registado o pico correspondente. A partir destes pontos, foi gerado um mapa através de uma técnica de interpolação geostatística (*kriging*) contendo a distribuição geográfica da doença ao longo dos meses em que ela ocorre em diferentes Estados (figura 6).

No que diz respeito às aplicações de análise das *doenças transmitidas por vectores*, estas representam uma parte muito significativa dos materiais publicados, sendo as metodologias semelhantes às utilizadas na análise dos factores físicos e

químicos, razão pela qual são aqui apenas brevemente referidos. Entre as doenças mais estudadas surge a malária, doença que atinge grandes proporções nos países em desenvolvimento. Hassan *et al* (2003) referem a construção de um SIG para a previsão do risco de malária no Egipto, processo que consistiu em desenvolver uma base de dados geográfica para a gestão de eventuais surtos e efectuar vários tipos de análises espaciais (essencialmente sobreposições de temas geográficos) e estatísticas (*análise de determinantes*).

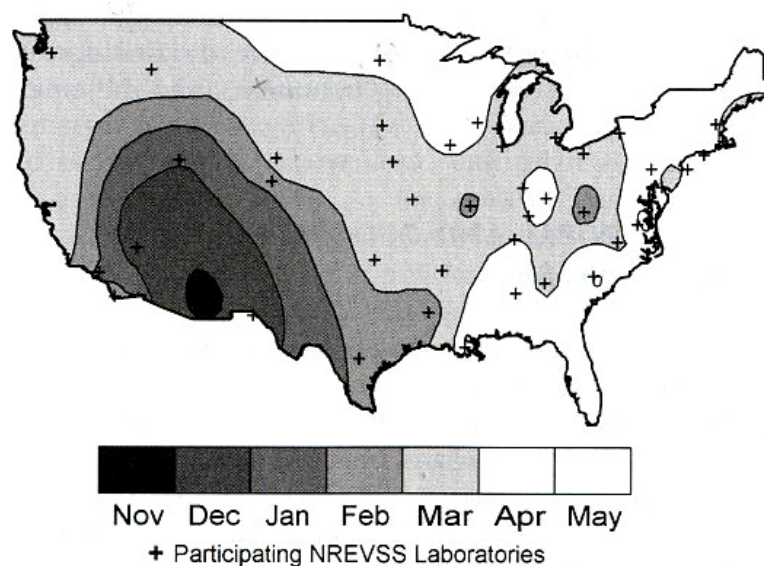


Figura 6 – Picos temporais e evolução das infecções por *rotavirus* nos EUA (Cromley e McLafferty, 2002).

Entre outros resultados, elaborou-se um mapa final identificando as regiões com maior risco. Os autores concluíram que o risco de malária pode ser previsto com base apenas em factores físicos: hidrogeologia e tipo de solo. Este estudo veio reforçar conclusões de estudos anteriores, que apontam os factores ambientais como intervindo de forma crucial na determinação do nível de risco das doenças transmitidas por vectores.

Existem ainda, no âmbito da geografia da doença, algumas aplicações de carácter mais particular. Boulos *et al* (2001) refere a informação espacial como toda aquela em que a localização tem alguma importância ou traz algum benefício, mesmo que a localização não se refira (na totalidade ou em parte) à localização geográfica na superfície terrestre. Muitas doenças e organismos tendem a escolher determinadas “localizações” anatómico-fisiológicas do corpo humano (sistemas ou órgãos do corpo) durante períodos temporais e de tratamento diversificados. Para

tirar partido desta situação, um exemplo consiste numa extensão (*BodyViewer*) para *ESRI ArcView 3.1*, que actua sobre registos clínicos contendo a localização geográfica (e.g. código postal) e também um código que refere a doença detectada, baseado na nomenclatura ICD-9 ou ICD-10 (Classificações Internacionais de Doença que codificam diagnósticos, complicações ou causas de morte), ligando estes registos a um tema vectorial contendo os sistemas/órgão referidos (figura 7).

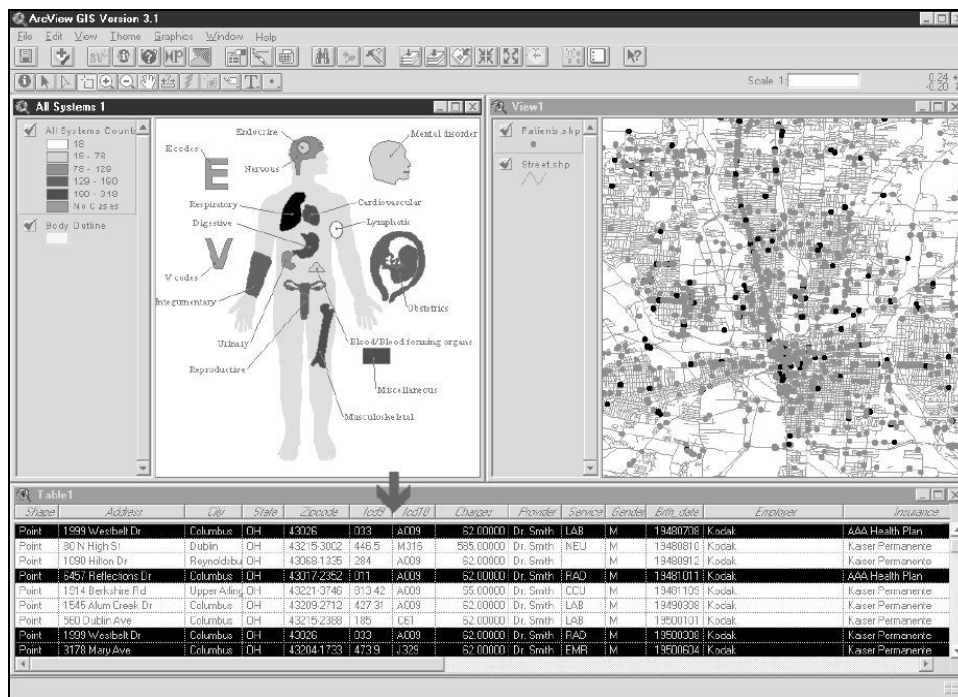


Figura 7 – Aspecto da aplicação *Body Viewer* (Boulos *et al*, 2001).

Este tema é também ligado a um tema geográfico, sendo possível visualizar onde os códigos ICD ocorrem, o que permite detectar, analisar e actuar eficazmente sobre novos padrões de distribuição de doenças. Por exemplo, as diferentes cores dos símbolos no tema que representa os órgãos do corpo humano representam a prevalência⁸ relativa de cada categoria de doença (e.g. doenças cardiovasculares) na base de dados geográfica dos pacientes, sendo que quanto mais escuro o símbolo, maior a prevalência. Na figura 7, foi seleccionado o símbolo de pulmão que representa todas as doenças respiratórias; todos os pacientes que possuem

⁸ Prevalência de uma doença consiste no número de casos existentes numa determinada população num determinado ponto do tempo. Beaglehole *et al*, (1993).

um código ICD de doença respiratória são simultaneamente identificados no tema geográfico e na tabela dos pacientes.

A tentativa de detecção de *clusters* constitui actualmente uma importante área de estudo ao nível das aplicações de saúde ambiental, mas não só, existindo várias outras aplicações para as técnicas existentes; a título de exemplo, Elliot e Wartenberg (2004) referem a importância da aplicação destes métodos no contexto dos sistemas de vigilância epidemiológica de doenças infecciosas.

O interesse que esta área tem suscitado levou ao desenvolvimento de várias técnicas, existindo já um número razoável de métodos de análise de *clusters* implementados nos actuais softwares de sistemas de informação geográfica. Estes métodos devem ser consideradas ferramentas exploratórias para ajudar nas decisões de investigadores e políticos, principalmente na análise de padrões espaciais complexos. A pesquisa de *clusters* espaciais não deve constituir um fim em si, mas sim uma ferramenta, entre outras, para o controlo e vigilância da Saúde Pública. Assim, quando um *cluster* é identificado, deve desenvolver-se uma investigação epidemiológica minuciosa para determinar se essa ocorrência é ocasional ou se está ligada a factores ambientais, sociais ou profissionais (ocupacionais). Cromley e McLafferty (2002) apresentam uma discussão muito completa das técnicas disponíveis. Segundo estes autores, a detecção de um *cluster* pode ocorrer tanto pela preocupação dos cidadãos, quando se apercebem da existência de um excesso de casos de uma determinada doença no ambiente em que se movimentam, ou quando o alarme é dado por um sistema de vigilância que procura sistematicamente padrões em dados recolhidos rotineiramente. A utilização dos SIG neste domínio destina-se a realizar análises confirmatórias, no sentido de comprovar a existência de um determinado *cluster*, ou exploratórias, quando procuram padrões espaciais deste tipo.

Algumas questões fundamentais devem ser sempre abordadas, independentemente do método utilizado. Em primeiro lugar, o número de ocorrências de uma doença deve ser sempre ponderado pela população em risco (conjunto de pessoas susceptíveis de contrair a doença em estudo). Estabelecer a escala de análise é também um factor especialmente crucial, na medida em que o mesmo fenómeno, analisado a escalas diferentes, pode apresentar resultados diversos; um exemplo desta situação é retratado por Jacquez e Greiling (2003), no decurso da análise de *clusters* da incidência de três tipos de cancro, no Estado de Nova Iorque como um todo, e na região de Long Island em particular. Em terceiro

lugar, surge a necessidade de estabelecer um conjunto de critérios para determinar a presença de *clusters* em termos de número de ocorrências, ou seja, quantas ocorrências em excesso são precisas para se considerar um *cluster*. O limiar a partir do qual um *cluster* pode ser entendido como significativo é de difícil definição, sendo comum o recurso a critérios estatísticos que calculam a probabilidade do aparecimento de *clusters* numa determinada população se dever ao acaso, existindo vários métodos para este efeito (e.g. Distribuição de *Poisson* ou ainda os Métodos de Simulação de Monte Carlo). A procura de *clusters* pode incidir sobre qualquer tipo de objecto espacial (pontos, linhas ou áreas) embora a maioria das técnicas incida sobre a busca de agregações de pontos ou áreas, bem como a maioria dos trabalhos realizados até ao momento. Pela maior relevância que apresentam para esta dissertação (discutida no capítulo sexto), optou-se por debater aqui as técnicas de detecção de *clusters* de áreas, podendo as restantes técnicas ser consultadas em Cromley e McLafferty (2002). Desta forma, no que respeita à busca de *clusters* de áreas, a técnica mais simples consiste na construção de mapas coropletos de taxas de incidência ou prevalência da doença; esta é uma situação em que a escala de análise tem muita importância, pois numa área pequena, a fiabilidade pode ser muito baixa devido ao impacto que uma ocorrência a mais ou a menos pode ter em distribuições contendo valores de observações muito reduzidos, situação designada em estatística como o problema dos números pequenos (*small numbers problem*). Uma forma de ultrapassar este problema consiste em recorrer a mapas de probabilidade, cartografando a significância estatística das taxas em vez dos seus valores. Outra opção para a análise de *clusters* de áreas consiste em recorrer a técnicas baseadas em medidas locais de dependência espacial, tais como o cálculo da estatística G^* , também conhecida como *método de Getis e Ord*, ou a indicadores locais de autocorrelação espacial, conhecidos como métodos LISA. Estes métodos medem a associação entre um valor registado num determinado local e os valores das áreas adjacentes ou próximas. As técnicas desenvolvidas são especialmente úteis para encontrar *clusters* baseados em dados areais; assim, um *cluster* corresponde a uma região (enquanto conjunto de áreas) que apresenta taxas anormalmente elevadas, ou seja, uma concentração local de taxas elevadas.

2.5.2 Geografia dos Sistemas de Cuidados de Saúde

Segundo Boulos *et al*, (2001); “A Geografia dos Sistemas de Cuidados de Saúde [...] lida com o planeamento, gestão e prestação de serviços de saúde adequados, garantindo, entre outras coisas, o adequado acesso dos pacientes a estes serviços, após determinar as necessidades da comunidade-alvo em termos de cuidados de saúde e as áreas de influência dos locais onde são prestados. Estes serviços, incluem também as actividades de prevenção e promoção da saúde”.

Em primeiro lugar, impõe-se definir *serviços de saúde*. Segundo Cromley e McLafferty (2002), são todos aqueles que têm como objectivo contribuir para melhorar a saúde das populações, subdividindo-se em *informais* (quando prestados pela família ou comunidade a que a pessoa pertence em ambiente familiar ou comunitário) ou *formais* (quando prestados por instituições públicas, privadas ou voluntárias através de hospitais, médicos, etc.). As contribuições encontradas na literatura referem em geral os serviços de saúde formais. Os autores analisam estas questões sob dois pontos de vista: a *análise do acesso aos serviços de saúde* e a *localização de serviços de saúde*. Albert *et al* (2000) abordam também a *análise da localização e distribuição espacial dos profissionais de medicina* e a *vigilância, monitorização e resposta a situações de emergência*. Os autores abordam esta temática de forma exaustiva, pelo que se optou por referir apenas algumas questões relacionadas com a *análise do acesso aos serviços de saúde*. Este é um conceito muito abrangente, relacionado com a capacidade que as pessoas têm de utilizar os serviços de saúde, quando e onde são necessários (Cromley e McLafferty, 2002). Existem cerca de cinco dimensões para o termo, mas apenas uma configura a dimensão geográfica: a *acessibilidade*. Este conceito implica barreiras geográficas, como a distância, os meios de transporte, o tempo de viagem e o custo, comparando a localização geográfica dos serviços com a da população que a eles recorre. Os autores referem que a utilização dos serviços de saúde tem a ver, antes de mais, com a *distância*: em geral, à medida que aumenta, a procura dos serviços de saúde diminui, embora existam algumas excepções. Uma questão frequentemente abordada nestes estudos é a determinação das áreas de influência dos hospitais e centros de saúde, que dizem respeito às áreas geográficas contendo a maioria da população servida pelas infra-estruturas. Albert *et al* (2000) referem a aplicação deste conceito num estudo levado a cabo na África do Sul, com o objectivo de analisar os efeitos do fim do *apartheid* no acesso aos hospitais da

província de *Natal/KwaZulu* em 1985. Este estudo recorreu ao software ESRI ARC/INFO para gerar *polígonos de thiessen* representativos das áreas de influência dos hospitais. Foram produzidos três mapas, cada um contendo diferentes rácios de pessoas por camas de hospital, tendo em conta as unidades reservadas à população branca, negra e população em geral (correspondente à remoção das barreiras raciais). Concluiu-se que o acesso da população negra a este serviço não se alterou de forma substancial com o fim do *apartheid*.

2.6 Conclusões

Embora tentando tipificar as principais aplicações e casos de estudo das tecnologias de informação geográfica na saúde, abarca-se aqui apenas uma pequena parte das possibilidades contidas na literatura, sendo o potencial de aplicação enorme. A expansão da literatura dedicada, aliada à disponibilização de ferramentas cada vez mais simples de utilizar, tem um papel importante nesta situação. Também em alguns países desenvolvidos, a aposta na construção de infra-estruturas nacionais de informação orientadas para a saúde integrando uma componente de informação geográfica tem um efeito impulsionador, tendo como objectivo ajudar a ultrapassar muitas das barreiras relacionadas com os dados, não apenas na sua disponibilização, mas também no que diz respeito à interoperabilidade entre formatos e sistemas.

A aposta na formação dos profissionais de saúde e/ou daqueles que com eles interagem é fundamental; só assim será possível ultrapassar os obstáculos relacionados com a motivação e o grau de conhecimento dos utilizadores que ainda se fazem sentir.

A maior parte dos desenvolvimentos citados neste capítulo foram concretizados nos EUA, sendo também de notar, na América do Norte, a importância do Canadá, primeiro país a possuir um SIG, como demonstrado pelos estudos de Shuai *et al* (2006) e de Blanton *et al* (2006) no campo dos sistemas de vigilância epidemiológica dos vírus do Nilo Ocidental e da raiva, respectivamente. Outros países desenvolvidos, especialmente no continente europeu, apresentam também bastante dinâmica. Boulos (2005) estudou os padrões de distribuição geográfica dos autores de publicações do periódico *Medical Informatics & the Internet in Medicine*, abrangendo contribuições de autores de 24 países, em quatro continentes, entre os anos de 1999 e 2004, concluindo que quase todos os artigos

tinham tido origem na Europa (81.25%), não se notando contribuições da África ou América do Sul. Não obstante, podem ser consultados, ao nível de outros periódicos (e.g. *International Journal of Health Geographics*), alguns estudos localizados nestes continentes. Todavia, grande parte destes estudos são efectuados por autores oriundos de países desenvolvidos de carácter académico e/ou são resultantes da aplicação de fundos de auxílio internacional (e.g. Rakotomanana *et al*, 2007, na prevenção e mitigação da malária nas terras altas de Madagáscar com recurso a Detecção Remota e SIG), constatando-se uma utilização incipiente ao nível dos sistemas nacionais de saúde destes países (e.g. na aposta sistemática em sistemas de vigilância epidemiológica que façam uso de ferramentas SIG).

Para além da geografia dos recursos, a geografia das doenças dita muitas das prioridades. Assim, a maior parte dos estudos no âmbito das doenças transmitidas por vectores (tais como a malária) ocorre em países em vias de desenvolvimento, enquanto a investigação em doenças causadas presumivelmente por causas ambientais (tais como o cancro) e na qualidade da prestação de serviços de saúde predomina geralmente nos países desenvolvidos, bem como a maior parte dos sistemas de vigilância epidemiológica em funcionamento. Por outro lado, algumas preocupações emergentes começam a mobilizar as autoridades de saúde, levando ao estabelecimento de sistemas de vigilância epidemiológica à escala global, focados em algumas doenças contagiosas graves, com potencial para originar uma pandemia, como a gripe aviária e a síndrome respiratória aguda grave (SARS). Esta última doença merece especial atenção, pela epidemia que teve lugar no sudoeste asiático, tendo atingido o seu pico na primeira metade do ano de 2003, sendo de realçar a importância que teve a utilização de informação geográfica no acompanhamento e gestão desta situação, como pode ser observado em Boulos (2004b).

No que diz respeito a Portugal, começa a notar-se o investimento do Serviço Nacional de Saúde (SNS) neste tipo de aplicações, como é exemplo o SESP, (apresentado mais adiante), mas não só. Alguns grupos de investigação ligados a universidades mostram já interesse nestes estudos, como é exemplo o Grupo de Análise de Sistemas Ambientais (GASA) da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, citando-se os trabalhos de Lopes *et al* (2005) e Alves *et al*, (2001), focados na avaliação do risco de infecção pela malária em algumas áreas de Portugal Continental através de técnicas de Detecção Remota e SIG.

Também as faculdades de Medicina se mostram interessadas, especialmente aquelas que desenvolvem actividades ligadas à Epidemiologia e Medicina Preventiva, sendo apresentado na secção 5.5.2 desta dissertação um estudo resultante da colaboração entre a Faculdade de Medicina da Universidade de Lisboa e a DGS. Existem também já volumes editados por autores portugueses, como é exemplo o estudo de Santana (2005), acerca das determinantes e distribuição espacial dos padrões da Saúde em Portugal. A recente realização, em 31 de Maio de 2007, do 1º Congresso de Saúde e Tecnologias de Informação Geográfica (STIG), parece demonstrar que existe em Portugal um interesse crescente por estas aplicações.

3 SISTEMAS DE VIGILÂNCIA EPIDEMIOLÓGICA E INFORMAÇÃO GEOGRÁFICA

3.1 Introdução

Como já foi aqui afirmado, os sistemas de vigilância epidemiológica assumem um papel de crescente importância na actual conjuntura global, por constituírem cada vez mais a primeira linha de defesa da Saúde Pública face às potenciais ameaças que enfrenta, impondo-se clarificar, antes de mais, em que consiste um sistema desta natureza. Por outro lado, existem já algumas implementações destes sistemas, com diferentes objectivos e graus de complexidade. Este capítulo define os sistemas e apresenta alguns exemplos de aplicação.

3.2 Sistemas de Vigilância Epidemiológica e Informação Geográfica

Importa, em primeiro lugar, apresentar o significado da actividade de vigilância epidemiológica no sentido de enquadrar o sistema que aqui se apresenta. Assim, em Lawson (2006), define-se e contextualiza-se a actividade de vigilância em Saúde Pública, como a recolha sistemática e contínua, análise e interpretação de dados de saúde com vista ao planeamento, implementação e avaliação das práticas de Saúde Pública, acompanhadas da disseminação em tempo útil destes dados aos decisores, com o objectivo último da sua aplicação na prevenção e controlo. Esta actividade é operacionalizada em sistema dedicados, usualmente designados por sistemas de vigilância epidemiológica. Nos nossos dias, a vigilância epidemiológica materializa-se numa multiplicidade de sistemas que actuam a diferentes escalas espaciais e temporais, e com diferentes graus de prioridade.

A tarefa prioritária destes sistemas prende-se com a vigilância de epidemias⁹. Relacionado com este conceito surge uma área particular da vigilância epidemiológica, que pode ser designada como vigilância de síndromes¹⁰ (*syndromic surveillance*). Embora em Lawson (2006), se associe este termo à recolha sistemática e contínua, análise e interpretação de dados que precedem o

⁹ Uma epidemia consiste na ocorrência, numa determinada comunidade ou região, de um número de casos que pode ser considerado anormalmente elevado ou inesperado, para a área ou para a época em causa. Beaglehole *et al* (1993).

¹⁰ Síndrome consiste num grupo de sintomas ou outras alterações nas funções corporais que, quando analisadas em conjunto, indicam a presença de uma determinada doença. Colin (2005).

diagnóstico (e.g. registos de ambulâncias, da aquisição de medicamentos prescritos, absentismo escolar ou ao emprego, sinais e sintomas recolhidos em consultas em situação de doença aguda, entre outros), com o objectivo específico de tentar detectar, a partir destes dados, um grau de probabilidade de ocorrência de um surto ou epidemia que possa ser considerado suficiente para despoletar uma investigação/intervenção das autoridades de Saúde Pública, este processo pode também ser designado como vigilância das “consequências”. Entre estas “consequências” estará a utilização dos serviços de saúde ou a monitorização do impacto social da doença enquanto *proxy* (indicador indirecto) da situação epidemiológica. Li *et al* (2006) definem, por seu lado, a vigilância de síndromes como a utilização de dados de saúde em bruto para a célere detecção de surtos de doenças ou ataques de bioterrorismo, embora, de forma mais precisa, a vigilância de síndromes se baseie na detecção de um conjunto de sinais e sintomas associados. Estes autores referem ainda que a disponibilidade de informação demográfica adicional (e.g. idade, sexo, localização) tem um papel muito importante na compreensão da etiologia e da dinâmica de um surto.

Um dos factores mais importantes a ter em conta em qualquer investigação de carácter epidemiológico tem a ver com a localização dos diversos fenómenos nela envolvidos. Como já citado, o primeiro estudo epidemiológico a fazer explicitamente uso da localização foi a cartografia dos óbitos por cólera, da autoria de John Snow, realizada em Londres, durante a epidemia de cólera de 1854. A sua intervenção pôs em evidência o facto de que, na maioria das situações, o estudo da ocorrência e propagação de doenças infecto-contagiosas requer não só uma análise de dados de saúde alfanuméricos, mas também uma análise espaço-temporal da sua relação com inúmeras variáveis ambientais.

Li *et al* (2006) discutem também a importância de dois tipos de vigilância epidemiológica distintos, mas inter-relacionados, aos quais estes sistemas devem poder dar resposta: por um lado, o seu uso na monitorização de rotina em Saúde Pública e, por outro, apoiar (e se possível antecipar) a tomada de decisão em situações de vigilância de síndromes, no decurso das quais tudo decorre geralmente de forma rápida. Todavia, o grau de preparação para estas situações está fortemente relacionado com a primeira fase, durante a qual se realiza investigação “de fundo”, acumula conhecimento e se prepara e planeia a intervenção em situações atípicas. No caso da monitorização de rotina, concretiza-se, em geral, na recolha e compilação de estatísticas sobre a ocorrência de um

grande número de patologias, as quais são utilizadas para dar suporte a vários estudos epidemiológicos de fundo. Assim, quando efectuada *per si*, a monitorização de rotina incide, na maior parte dos casos, sobre patologias cuja evolução e prevalência numa determinada população não apresenta normalmente um risco de aumento repentino, ou seja, cujo risco para a Saúde Pública não apresenta variações repentinas ou de grande magnitude, sendo exemplos o cancro, a obesidade e doenças relacionadas, e as doenças não transmissíveis em geral. Quanto às situações de vigilância de síndromes, focam-se, pelo contrário, em doenças cuja prevalência pode aumentar subitamente, podendo acarretar graves riscos para a Saúde Pública se não forem tomadas medidas de contenção/mitigação no sentido de evitar a sua escalada, mas são, em geral, precedidas e/ou acompanhadas por monitorização de rotina.

Finalmente, de acordo com os desenvolvimentos mais recentes, é possível afirmar que, o estado da arte se encontra nos sistemas de vigilância epidemiológica multiplataforma que integram várias tecnologias, entre as quais se incluem as tradicionais bases de dados relacionais, sistemas de informação geográfica e mecanismos de disponibilização de informação na Internet.

3.3 Domínios e escalas de intervenção

Como já afirmado, estes sistemas intervêm tanto na monitorização de rotina em Saúde Pública, como na tomada de decisão em situações críticas. Em ambos os casos, as iniciativas e instituições intervenientes repartem-se entre o âmbito global, regional ou nacional. Ao nível global, destaca-se a *Organização Mundial de Saúde* (OMS), parte integrante das Nações Unidas, no âmbito regional aponta-se o recente esforço de coordenação na União Europeia, com a criação do *European Centre for Disease Prevention and Control* (ECDC), e ao nível nacional surgem, na vanguarda da vigilância epidemiológica, os *Centers for Disease Control* (CDC) dos Estados Unidos da América e a *Public Health Agency of Canada*. Estes organismos constituem referências para estudos na área, na medida em que implementam e controlam sistemas de vigilância epidemiológica de vários tipos, coordenam esforços e comunicação entre actores, levam a cabo intervenções que exigem a mobilização de meios, por vezes maciça, e emitem directivas não só para a implementação destes sistemas, mas também para a sua avaliação, como pode ser consultado numa área específica do sítio na Internet do CDC (CDC, 2001). As

organizações já referidas nos vários âmbitos territoriais têm também como responsabilidade definir as doenças que se encontram sujeitas a vigilância, o que condiciona, por sua vez, o surgimento de diferentes sistemas e ferramentas. Por exemplo, a *Public Health Agency of Canada* (Health Canada, 2000) publica uma lista de doenças transmissíveis sobre as quais deve ser exercida vigilância nacional do tipo *syndromic surveillance*, que inclui seis tipos principais:

- Doenças entéricas¹¹, presentes nos alimentos e na água;
- Doenças transmitidas por contacto directo e pelas vias respiratórias;
- Agentes patogénicos¹² transmitidos pelo sangue e contacto sexual.
- Doenças transmitidas por vectores e outros agentes zoonóticos¹³;
- Doenças prevenidas pela vacinação de rotina;
- Potenciais agentes de bioterrorismo à escala mundial.

3.4 Exemplos de aplicação

Nos nossos dias, estes sistemas encontram-se por vezes fragmentados entre as escalas de análise global, regional, nacional e local, não existindo ainda uma integração e articulação eficaz, como seria desejável, no sentido de aumentar a sua eficiência através do funcionamento em rede. Todavia, o panorama é de optimismo, pois a tendência futura será a integração de vários sistemas actualmente desconexos, estando já em curso, na União Europeia, através do Projecto *TESSy*, a estruturação de um sistema integrado europeu de vigilância de doenças transmissíveis (Balkányi, 2007). Outro exemplo é o desenvolvimento da *Canadian Network for Public Health Intelligence (CNPHI)*, que permite congrega e analisar dados de vários sistemas regionais, como por exemplo, na vigilância do Vírus do Nilo Ocidental (*West Nile Virus*) neste país (Gosselin *et al*, 2005). Entre os vários casos de estudo documentados, são aqui apresentados três exemplos ilustrativos destes sistemas, nas escalas de análise global e nacional. Pretende-se, através dos exemplos apresentados, ilustrar a infra-estrutura, os dados que utilizam, as análises que efectuam e os produtos (indicadores) que fornecem.

¹¹ Infecções bacterianas ou virais do tracto intestinal. NIAID (2008).

¹² Microorganismo capaz de originar uma doença. Colin (2005).

¹³ Agentes responsáveis por doenças dos animais transmissíveis ao homem (*zoonoses*). Colin (2005).

3.4.1 Vigilância global da Gripe Aviária Altamente Patogénica

Aproximadamente a partir de Janeiro de 2004, tornou-se perceptível que o vírus H5N1, causador da Gripe Aviária Altamente Patogénica (*HPAI - Highly Pathogenic Avian Influenza*), em aves selvagens e domésticas, tinha ultrapassado a barreira da espécie e passado a ser transmissível aos seres humanos (WHO, 2005). Os primeiros casos surgiram no Vietname, Tailândia e Camboja, mas o facto das aves migratórias actuarem como um dos vectores da doença fez reacear um cenário de pandemia, com a disseminação desta doença a nível global. Embora tal cenário não se tenha, até à data, concretizado, esta é uma doença sujeita, desde então, a vigilância constante a vários níveis. À escala global, tinha sido já iniciado pela FAO (*Food and Agriculture Organization*) desde 1994, o programa EMPRES (*Emergency Prevention System for Transboundary Animal and Plant Pests and Diseases*), com o objectivo de monitorizar e controlar doenças animais cuja difusão pode facilmente ultrapassar as fronteiras nacionais (Martin *et al*, 2007). Um dos componentes gerado no âmbito deste programa consiste num sistema de vigilância epidemiológica capaz de gerar alertas precoces, designado *EMPRES-i (Global Animal Health Information System)*. É constituído por vários módulos, e integra, entre outras, ferramentas de informação geográfica, sendo no essencial uma plataforma de partilha de informação entre os responsáveis da área de saúde animal da FAO e os peritos e instituições envolvidos na gestão de surtos e resposta a emergências. Tecnicamente, o núcleo deste sistema apoia-se numa base de dados *Oracle*, acessível pela Internet, que concentra dados relativos aos casos ocorridos de várias doenças animais cuja vigilância é considerada prioritária, entre as quais a gripe aviária altamente patogénica, febre aftosa (*foot and mouth disease*), febre do vale do rift (*rift fever valley*) e a peste suína africana.

Com o início da crise relacionada com a gripe aviária em 2004, o sistema assumiu uma nova importância, uma vez que uma das suas tarefas prioritárias passou a ser o registo das ocorrências desta doença em aves domésticas e selvagens. Os dados são recolhidos a partir de uma multiplicidade de fontes, entre as quais a OMS, a Organização Mundial para a Saúde Animal, a Comissão Europeia e os governos dos países aderentes. Os produtos disponibilizados aos decisores consistem, em primeiro lugar, em relatórios de situação, tais como a Lista de Acompanhamento de Doenças (*Disease Tracking List*), que fornece um sumário bissemanal da situação epidemiológica mundial da gripe aviária em termos da sua

distribuição espacial e temporal, sendo evidenciados todos os surtos confirmados e investigações pendentes durante as duas semanas anteriores. Outro produto, um pouco mais elaborado, é o Ponto de Situação da Gripe Aviária do Centro para as Doenças Animais Transfronteiriças (*Emergency Centre for Transboundary Animal Disease HPAI Situation Update*), que reúne informação relativa a surtos, suspeitas de surtos ou outros assuntos relevantes nos últimos dois meses, tendo a particularidade de apresentar na primeira página um mapa que ilustra a distribuição geográfica dos surtos confirmados de gripe aviária em aves domésticas ou os casos de aves selvagens encontradas mortas. Numa escala temporal mais reduzida, o Relatório de Actualização da Gripe Aviária Altamente Patogénica (*HPAI Update*) é produzido e distribuído em formato electrónico três vezes por semana (Segunda, Quarta e Sexta-Feira), mas a sua distribuição é restrita, uma vez que inclui também suspeitas e rumores. Para colmatar esta limitação, a FAO emite mensalmente as Noticias de Emergências de Gripe Aviária (*Avian Influenza Disease Emergency News*), destinado à divulgação pública, incluindo uma curta análise da situação, o registo das situações de surto confirmadas a nível nacional, e ainda informações sobre missões e projectos da FAO. Alguma desta informação, dependendo da sua relevância para o público-alvo e grau de sensibilidade, é ainda publicada no sítio na Internet dedicado do sistema EMPRES (FAO, 2008). Para além dos relatórios que produz, e que constituem a sua face mais visível, este sistema constitui uma ferramenta de análise, uma vez que a sua base de dados central (*Oracle*) se encontra ligada a um sistema de informação geográfica, permitindo não só efectuar interrogações convencionais à base de dados, mas também realizar uma série de operações de análise espacial que permitem visualizar os padrões de distribuição espacial da doença e explorar as relações entre localização, factores ambientais e doença, contribuindo assim para a compreensão dos factores epidemiológicos e ecológicos responsáveis pela eclosão e alastramento das doenças animais transfronteiriças. As operações que este sistema permite levar a cabo sobre a informação geográfica apresentam vários tipos e graus de complexidade. Os dados de vigilância introduzidos no EMPRES-i, relativos à localização dos casos ocorridos, podem ser visualizados sob a forma de simples cartografias temáticas (figura 8), sendo depois possível cruzá-los com informação auxiliar, tal como uso/ocupação do solo, densidade de aves domésticas presentes no território, imagens de Detecção Remota ou outras variáveis que possam ser relevantes para a análise da epidemiologia da doença.

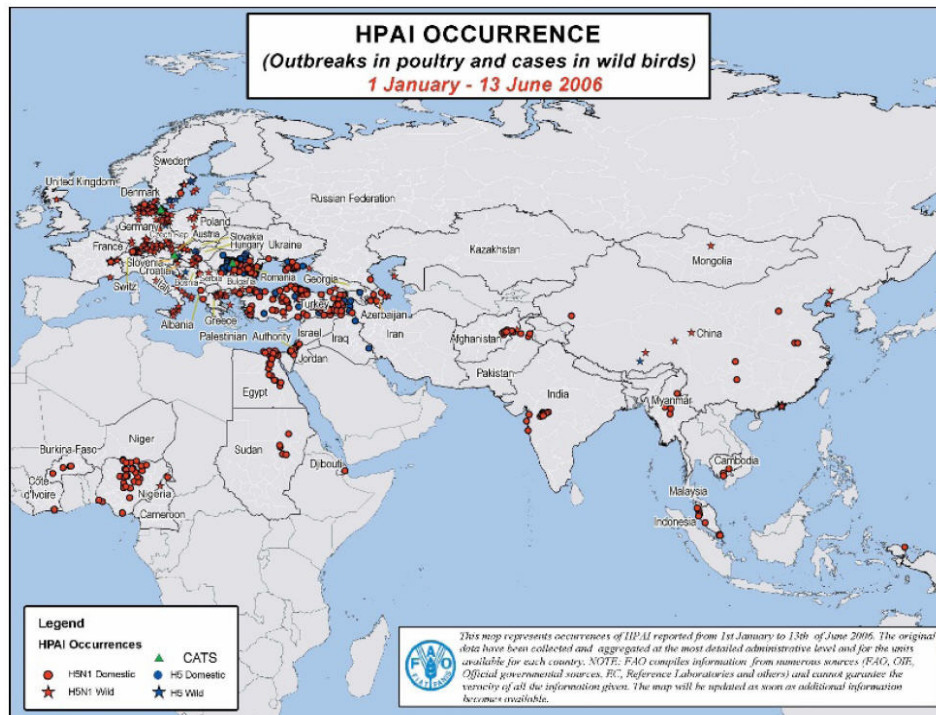


Figura 8 – Surtos de gripe aviária altamente patogénica entre Janeiro e Junho de 2006 (Martin *et al*, 2007).

Outro tipo de ferramentas suportadas por este sistema, e relevantes em epidemiologia espacial, são as operações de análise espacial exploratória, as quais permitem detectar o aparecimento de *clusters* de casos ou estabelecer associações entre a magnitude da doença e outros factores de risco. Um exemplo (figura 9) foi a preparação, em 2006, de um mapa representando a densidade de ocorrências, decorrente da aplicação, a cada localização, de uma função do tipo *kernel*, a qual consiste numa “janela” móvel com determinadas dimensões, que calcula um rácio entre o número total de surtos localizados dentro dos seus limites e a respectiva área. Esta cartografia suporta a identificação de áreas onde a densidade de surtos é máxima, e no caso em estudo, permitiu observar que a maior quantidade de animais infectados pareceu ter sofrido um desvio do sudoeste asiático para as zonas central e ocidental da Europa, no período em estudo.

Uma vez encontrados *clusters*, uma análise possível consiste em avaliar a sua proximidade a possíveis factores de risco ambiental relacionados com estes surtos, tais como zonas húmidas e arrozais, cuja delimitação, na maior parte das vezes através de Detecção Remota, tem sido objecto de outros estudos (Gilbert, 2005).

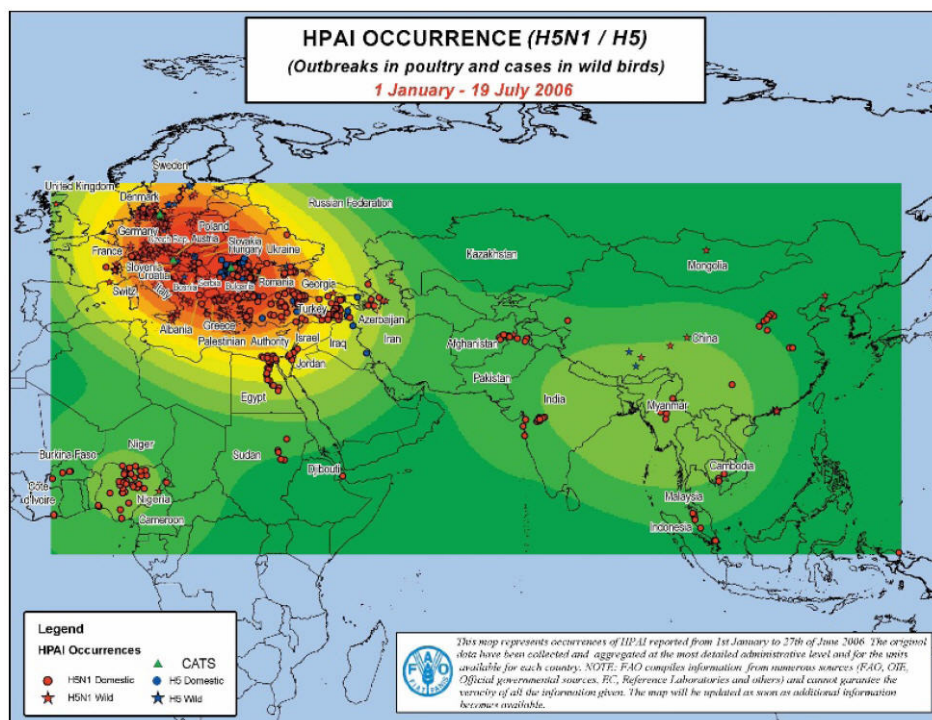


Figura 9 – Densidade de surtos de Gripe Aviária Altamente Patogénica (Martin *et al*, 2007).

Outra análise possível consiste na utilização de modelos digitais de terreno para obter visualizações tridimensionais do território no sentido de identificar rotas preferenciais para a difusão da doença, uma vez que, pelo menos no que respeita às aves domésticas, os comerciantes se deslocam pelos vales, o que é especialmente válido em países com áreas de grande altitude e/ou desnível, como é o caso da fronteira entre Paquistão e Afeganistão. São ainda disponibilizadas, no sítio da Internet do EMPRES-i, algumas ferramentas destinadas a profissionais de Epidemiologia e Veterinária que podem ser descarregadas e instaladas em computadores pessoais, permitindo modelar espacialmente fenómenos relacionados com a doença (*TADinfo* e *TADSimulator*).

3.4.2 Vigilância do Vírus do Nilo Ocidental na província do Quebec, Canadá

Apesar de existirem indícios da ocorrência de infecções pelo vírus do Nilo Ocidental (*West Nile Virus - WNV*) na região do Médio Oriente durante séculos, este vírus só foi detectado na América do Norte em 1999, mais concretamente nos Estados Unidos da América, Estado de Nova Iorque, tendo o número de infecções aumentado dramaticamente no ano de 2002 e atingido o auge no ano de 2003, com

9862 casos confirmados neste país, dos quais 264 se revelaram fatais (Gosselin *et al*, 2005). Mais a norte, no Canadá, a província do Quebec registou, em 2002, 20 casos, com três óbitos. Esta é uma doença transmitida por vector, neste caso por mosquitos, podendo infectar o Homem, mas também algumas espécies animais, tais como várias espécies de *corvidae* (corvos), cavalos e potencialmente outros animais domésticos. A perspectiva de escalada desta situação levou o governo regional a adoptar um plano de protecção da Saúde Pública específico para esta doença, visando prevenir a ocorrência de casos humanos, detectar precocemente a presença deste vírus numa determinada área, identificar áreas de risco de transmissão ao Homem para efectuar acções de prevenção e monitorização, e determinar o nível de transmissão humana. Entre as áreas de intervenção prioritárias deste plano e níveis de desempenho pretendidos, encontrava-se uma componente de monitorização, materializada num sistema de vigilância epidemiológica integrado (combinando a monitorização de casos ocorridos em humanos, aves, mosquitos e outros animais) e capaz de receber e processar dados em tempo real.

Para dar resposta a estas especificações, o *Institut National de Santé Publique du Quebec* (INSPQ) montou, em 2003, o sistema *ISPHM-WNV (Integrated System for Public Health Monitoring of the West Nile Virus)*, com o objectivo de fornecer aos decisores ferramentas para avaliar e recomendar intervenções de forma simples e rápida. A informação obtida ao nível desta província converge, tal como no caso de outros sistemas regionais, para uma rede nacional (*Canadian Network for Public Health Intelligence - CNPHI*), permitindo a coordenação a várias escalas de análise.

O sistema *ISPHM-WNV* recebe dados de muitas fontes, introduzidos tanto por entidades oficiais de Saúde Pública (a província do Quebec encontra-se dividida em dezoito autoridades de saúde, responsáveis pela notificação obrigatória de todos os casos de infecção humana por este vírus), como pela população, naquilo que consiste um exemplo de participação da sociedade a vários níveis numa causa comum. Assim, foi feito um apelo à população em geral para reportar o aparecimento e localização de corvos mortos (a ocorrência de *clusters* destes casos pode indicar um potencial foco de amplificação na transmissão da doença) através de um centro telefónico, designado *WNV-Info line*. A monitorização dos mosquitos é efectuada pelas autoridades competentes em estações permanentes e também, sempre que necessário, em áreas de risco assinaladas pela detecção de *clusters* de aves infectadas. Um painel de peritos interage ainda com os decisores

em Saúde Pública, com os administradores do sistema, e com os responsáveis pela introdução local dos dados, para ajudar na interpretação de resultados e tomada de decisões. A entrada neste sistema é feita através de um sítio da Internet dedicado, de acesso restrito, que funciona como plataforma de realização das análises e troca de documentos entre intervenientes, permitindo também a actualização diária da base de dados de monitorização. A arquitectura do sistema baseia-se em três componentes principais: uma base de dados centralizada, desenvolvida em *Microsoft SQL Server*, que armazena todos os dados (alfanuméricos e espaciais); software *WebGIS*¹⁴, cuja solução consiste numa suite de ferramentas *Open Source*, designada *JMap*, baseada na plataforma *Java*, que suporta as funcionalidades de análise e manuseamento de informação geográfica; estrutura do sítio da Internet propriamente dita, que aloja e organiza todas as funcionalidades do *ISPHM-WNV*. Uma vez feito o acesso, torna-se possível aceder aos resultados e/ou realizar várias operações, das quais as mais importantes são a introdução da localização geográfica dos vários tipos de ocorrências definidas na base de dados, a realização de várias operações de análise espacial (desde representações cartográficas simples das ocorrências, até outras mais complexas, como a detecção de *clusters*) e a visualização de estatísticas sob a forma de gráficos e tabelas, ilustrando a frequência e evolução temporal dos vários eventos (casos de infecção ocorridos) armazenados na base de dados.

No que diz respeito às representações cartográficas, o sistema permite produzir mapas ilustrando os casos de infecções detectados pelas quatro redes de monitorização que o alimentam de dados (humanos, corvos, mosquitos, outros animais) sob forma pontual ou areal, agregados por uma divisão administrativa à escolha, e ainda combinar estes com informação auxiliar, tal como zonas de vegetação, hidrografia, redes de comunicação, entre muitos outros tipos de informação auxiliar armazenada. A figura 10 ilustra a área de Montreal em 2005, indicando para as diferentes subdivisões administrativas desta zona a quantidade reportada de corvos mortos, subdividindo os resultados (e.g. aves infectadas ou não, diagnóstico indeterminado) através do recurso a gráficos circulares. Quanto às funcionalidades mais avançadas, inclui uma *interface* preparada especificamente para detectar e representar cartograficamente *clusters* espaço-temporais de corvos

¹⁴ Por constituir a designação mais comum na literatura, optou-se por utilizar este termo de origem anglo-saxónica, referindo-se a um SIG cujas funções são disponibilizadas através de uma interface de acesso remoto, recorrendo à Internet como meio para esta disponibilização.

mortos, desenvolvida com base no método *SaTScan* (Kulldorff, 1997), para além de incluir ferramentas específicas para filtragem da informação contida nos mapas.

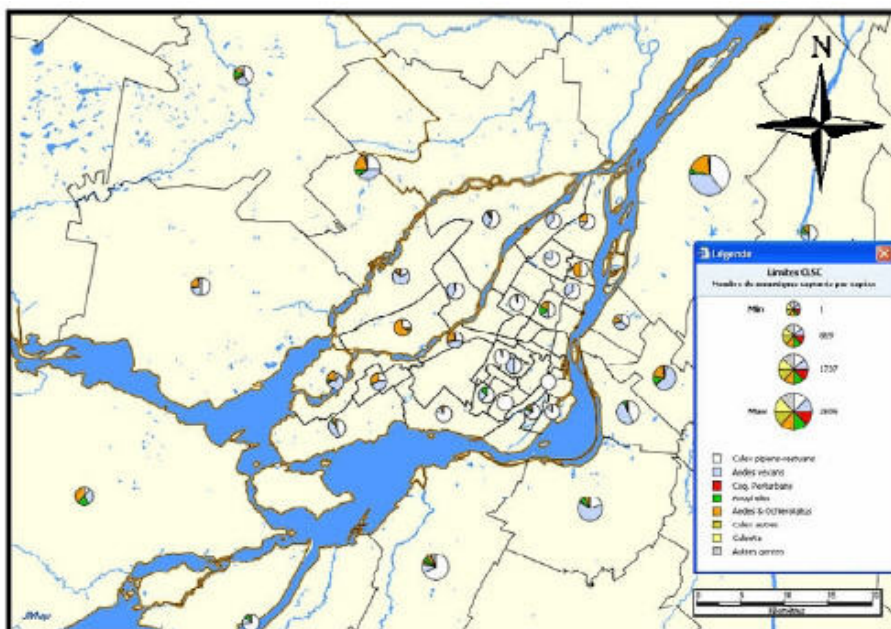


Figura 10 – Análise das recolhas de corvos mortos na área de Montreal em 2005 (Gosselin *et al*, 2005).

Durante a época do ano em que este vírus se considera potencialmente activo, a actualização do sistema e a análise de resultados são feitas diariamente, podendo ser propostos aos decisores diferentes níveis de intervenção, consoante uma tipologia de resultados obtidos: ausência de actividade, possível fonte de infecção (corvos mortos encontrados), provável fonte de infecção (corvos e mosquitos infectados) e fonte de infecção confirmada (casos de infecção detectados em humanos). As diversas acções a efectuar são planeadas através do recurso a uma grelha de análise multicritério. Os resultados são disseminados a vários níveis, desde uma esfera mais restrita, tal como a rede nacional (CNPFI) atrás referida, até ao conhecimento do público em geral, sendo para tal produzido semanalmente um mapa, divulgado através de um boletim electrónico. Um aspecto muito interessante do sistema e que constitui sem dúvida um exemplo de boas práticas é o facto do seu funcionamento ser avaliado anualmente após cada época de vigilância, por questionário destinado aos utilizadores, englobando cerca de 13 parâmetros qualitativos, sendo assim possível aperfeiçoar este sistema numa base anual.

3.4.3 Vigilância da Meningite Meningocócica na Alemanha

O sistema *EpiScanGIS* (Reinhardt *et al*, 2008) foi lançado em Maio de 2006 na Alemanha. É disponibilizado através de um sítio na Internet (*EpiScanGIS*, 2005) e destina-se à vigilância em tempo real dos casos de Meningite Meningocócica, combinando registos de ocorrências com informação demográfica, tal como idade, incidência e densidade populacional. Divide-se em duas áreas, uma das quais, de acesso público, fornece mapas gerados dinamicamente (figura 11) e actualizados semanalmente.

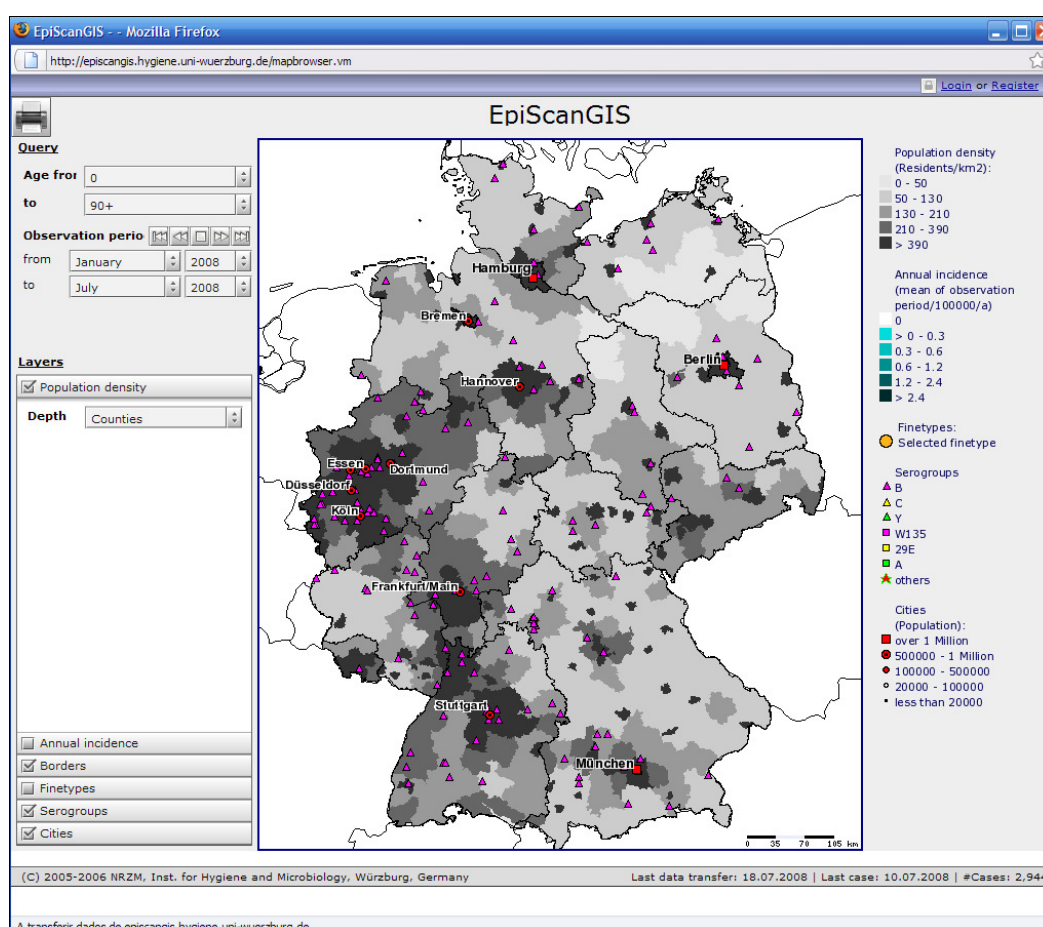


Figura 11 – Aspecto do *EpiScanGIS* (área pública) (*EpiScanGIS*, 2005).

Outra área, de acesso restrito aos decisores em Saúde Pública, contém um sistema de alerta precoce, baseado na detecção automática pelo sistema, de *clusters* espacio-temporais através do método *SaTScan*, os quais são também dispostos em mapas. De referir que este método é fornecido gratuitamente, através de aplicação

homónima, para instalação em computadores pessoais, requisitada a partir de um sítio na Internet (SaTScan, 2005). Outra das características do *EpiScanGIS* que constitui um rumo exemplar para os sistemas de vigilância em Saúde Pública, é a aposta em componentes de *software* do tipo *Open Source*, não-comercial. A sua arquitectura baseia-se nas aplicações *PostgreSQL* (base de dados objecto-relacional do sistema), *PostGIS* (suporta a realização de interrogações a objectos geográficos guardados na base de dados), *UMN Mapserver* (geração de mapas com base nos dados armazenados na base de dados) e *OpenLaszlo* (geração de uma aplicação, baseada em tecnologia *Flash*, que permite o acesso aos dados através da Internet).

3.5 Conclusões

Como fica ilustrado neste capítulo, os sistemas de vigilância epidemiológica são necessários e encontram-se cada vez mais presentes na estratégia de gestão e controlo da Saúde Pública de muitos países, sendo um exemplo flagrante de como se pode tirar partido do recurso a tecnologias recentes nesta área do conhecimento para reforçar mecanismos de vigilância do estado da Saúde Pública que eram, na sua maior parte, e até há relativamente pouco tempo, *retrospectivos* (exemplo da rede de Médicos Sentinela em Portugal, desenvolvida no capítulo seguinte) e neste momento, graças à implementação de sistemas baseados na transmissão de dados em tempo real e na vigilância de sintomas, prometem assumir um carácter *prospectivo*.

Outra característica essencial destes sistemas actuais reside, cada vez mais, na diversidade tecnológica que apresentam, constituindo, na sua maioria, como se verifica pelos exemplos atrás apresentados, plataformas de integração de tecnologias muito variadas (bases de dados relacionais, sistemas de informação geográfica, comunicações, entre outras). A aposta em ferramentas de *software* não comercial é outra faceta que começa a fazer-se sentir aqui, e desde que devidamente gerida, poderá trazer vantagens a vários níveis. Assim, a ausência de outros custos a não ser aqueles de desenvolvimento inicial das aplicações e respectiva manutenção, poderá facilitar a implementação destas ferramentas em países possuidores de escassos recursos, para além de libertar fundos para outros investimentos ao nível do orçamento que os países mais desenvolvidos empregam na gestão da saúde dos seus cidadãos.

Por último, a evolução no sentido da integração e compatibilização técnica dos sistemas já existentes em funcionamento (e dos vindouros) é uma necessidade, constituindo o Canadá um exemplo de boas práticas a seguir, mas existindo ainda exemplos no sentido inverso, como é o caso do sistema *RODS (Real-Time Outbreak and Disease Surveillance System)*, que apresenta soluções técnicas e metodológicas inovadoras, mas encontra-se operacional apenas nos estados norteamericanos da Pennsylvania e Utah (Boulos, 2004).

4 MODELAÇÃO DO SISTEMA DE SUPORTE A EMERGÊNCIAS EM SAÚDE PÚBLICA

4.1. Introdução

No contexto das actuais necessidades de vigilância em Saúde Pública, a DGS encontra-se a desenvolver uma infra-estrutura nacional para gestão de dados de Saúde Pública e apoio à decisão em situações de vigilância epidemiológica, implementada em ambiente multiplataforma e integrando, entre outras, uma componente de informação geográfica. Esta infra-estrutura materializa-se no SSESP, aqui apresentado e discutido sobre diversos ângulos de análise, desde a modelação conceptual até à sua implementação e estado da arte.

4.2 Estabelecimento do estado da arte e análise de benefícios

O passo inicial deste processo consistiu numa análise do estado da arte da vigilância epidemiológica efectuada pela DGS, o que levou os decisores em Saúde Pública da instituição a uma reflexão acerca dos benefícios de investir na implementação do sistema, implicando no essencial renovar algumas funcionalidades já existentes e implementar outras totalmente novas.

Anteriormente à implementação do SSESP, a vigilância epidemiológica da gripe (a perspectiva de uma pandemia de gripe constituiu o ponto de partida para a concretização do sistema) era levada a cabo pela DGS, com base nos seguintes recursos:

- Sistema de Informação das Urgências Hospitalares e Centros de Saúde (SIU);
- Rede de Médicos Sentinela.

No que diz respeito à Rede de Médicos Sentinela, que ainda se mantém, esta permite obter uma estimativa bastante boa da incidência e perfil epidemiológico das epidemias de gripe. Todavia, o seu tempo de resposta é demasiado lento (relatórios semanais) para sustentar o apoio à decisão ao nível das medidas de Saúde Pública em caso de pandemia. Para além disso, o facto da rede ser constituída por médicos voluntários leva a uma distribuição pelo país algo irregular, não cobrindo

sistematicamente o território e, conseqüentemente, que a informação proveniente desta rede se revele pouco representativa. Por outro lado, é discutível a sustentabilidade de esta rede em caso de pandemia, sendo provável que um número razoável destes profissionais possa vir a ser também afectado, comprometendo seriamente a integridade da rede ou a periodicidade dos relatórios.

Quanto ao SIU, este consistia numa base de dados relacional em formato *Oracle 9i* (actualmente substituída pela base de dados *Oracle 10g* do SESP), não espacial, armazenando dados relativos à procura de serviços de urgência de hospitais e centros de saúde, actualizada diariamente por transferência dos registos originados no acto da inscrição dos utentes quando recorrem a um serviço de atendimento em situação de doença aguda (vulgarmente designado como serviço de urgência) em qualquer hospital (sistema SONHO) ou centro de saúde (sistema SINUS). Por cada registo de ida às urgências em hospitais e centros de saúde do território continental português, eram também armazenados no SIU dados referentes ao estabelecimento de saúde de atendimento, idade e sexo do utente, causa da ida à urgência (e.g. traumatismo, doença), freguesia de residência, etc. No âmbito deste sistema, a DGS possuía ainda um sítio na Internet onde procedia à publicação, numa área de acesso restrito, de quadros e gráficos ilustrando informação obtida a partir do processamento dos dados armazenados, sendo a actualização deste sítio efectuada diariamente através de um conjunto sequencial de processos manuais realizados por um operador, e com recurso a aplicações de produtividade comuns (*Microsoft Access 2003*, *Microsoft Excel 2003* e *Microsoft Word 2003*). Para além da morosidade e propensão a erros manuais do processo de actualização, que exigia a alocação diária de um operador, o recurso a estas ferramentas possuía várias deficiências, entre as quais o facto do *Microsoft Access* possuir poucas funcionalidades enquanto Sistema Gestor de Bases de Dados (SGBD), apresentando por exemplo limitações de capacidade e desempenho no processamento diário de grandes volumes de dados, realizado neste ambiente através de um conjunto de rotinas (macro), com estabelecimento prévio de uma ligação *Open Database Connectivity* (ODBC) à base de dados *Oracle 9i*.

Após a avaliação inicial, decidiu-se que a estratégia geral para o futuro consistiria em gerar um novo sistema tendo como base o SIU. Após uma reflexão custo/benefício contemplando alterações e novas funcionalidades do sistema, vários pontos foram identificados como essenciais para a obtenção de um sistema

mais eficiente e produtivo para as tarefas e desafios da vigilância epidemiológica actual. Assim, impunha-se:

- Substituir todo o aparato manual de processamento diário de dados por um mecanismo inteiramente automático;
- Fazer evoluir a infra-estrutura existente para uma plataforma de integração, combinando o armazenamento e processamento simultâneo de dados alfanuméricos e geográficos, com todas as vantagens que estes últimos possuem por estarem associados a um determinado local do espaço, permitindo uma fácil e rápida compreensão dos padrões de distribuição espacial dos fenómenos, desde há muito reconhecida como um factor chave em Epidemiologia e no apoio à decisão em Saúde Pública;
- Centrar todas as tarefas de armazenamento e processamento de dados alfanuméricos numa aplicação robusta (e.g. *Oracle* ou *Microsoft SQL Server*) que possua a capacidade e arquitectura de um verdadeiro SGBD. Esta opção traz várias vantagens (Ramakrishnan e Gehrke, 2000), tais como: melhor garantia da independência dos dados, fundamental num sistema integrado, onde várias aplicações acedem à base de dados central; maior eficiência no acesso aos dados, obtendo-se também melhorias substanciais no processamento diário de grandes volumes de dados; manutenção de um bom nível de integridade e segurança, fundamental quando estes apresentam grandes volumes, na casa dos vários milhões de registos anuais; melhor administração de dados, permitindo, por exemplo agendamento transparente da execução das tarefas diárias;
- Renovar totalmente o sítio na Internet existente, tornando-o capaz de apresentar todos os resultados do sistema (alfanuméricos, gráficos e geográficos) diariamente actualizados, aumentar o seu grau de interacção com o utilizador (a estrutura existente era, na sua maior parte, estática) e a sua usabilidade (Nielsen e Loranger, 2006).

4.3 Definição do problema e objectivos

Antes de se dar início à construção desta infra-estrutura foi necessário definir quais as situações específicas a que deveria dar resposta, bem como as funcionalidades

necessárias para o efeito. Para suportar a definição de objectivos, foram consultadas as fontes bibliográficas disponíveis e os principais intervenientes (especialistas em Epidemiologia, autoridades em Saúde Pública e técnicos envolvidos na construção do sistema), tendo sido ainda avaliados os recursos existentes. Concluiu-se que, para dar resposta aos seus objectivos de vigilância epidemiológica, este sistema deve recorrer a várias fontes de dados e a uma arquitectura específica, fornecendo automaticamente, numa base diária e sob várias formas, vários indicadores, que devem poder ser consultadas no ambiente da DGS, mas também remotamente. Deve, nomeadamente, providenciar:

- A procura diária de urgências hospitalares, ambos os sexos, em Portugal Continental e por Distrito, para todas as idades e também por classe etária (0-14, 15-24, 25-44, 45-64 e maiores de 65 anos de idade);
- A procura diária de urgências em centros de saúde, ambos os sexos, em Portugal Continental e por Distrito, todas as idades e também por classe etária (0-14, 15-24, 25-44, 45-64 e maiores de 65 anos de idade);
- A procura diária de urgências (hospitais + centros de saúde), ambos os sexos, em Portugal Continental e por Distrito, todas as idades e também por classe etária (0-14, 15-24, 25-44, 45-64 e maiores de 65 anos de idade).

As diversas formas de apresentar a resposta às interrogações efectuadas ao sistema correspondem aos produtos do sistema, tendo-se optado por produzir gráficos de linhas, relatórios e mapas temáticos que permitem monitorizar a evolução diária da procura nas urgências. Esta opção combinou a experiência acumulada no âmbito do SIU com as pesquisas efectuadas, existindo já sistemas com arquitecturas semelhantes, tais como o *ISPHM-WNV* (Gosselin *et al*, 2005) e o *GeoMedStat* (Li *et al*, 2006). Decidiu-se ainda pela inclusão de análises do nível de alerta, divididas entre a monitorização gráfica da situação semanal e a emissão diária de avisos através de correio electrónico, ambas alimentadas pela comparação com valores epidemiológicos de referência (*baselines*); esta questão será desenvolvida no capítulo quinto.

Por outro lado, para além da elaboração diária dos produtos atrás referidos, este sistema deve permitir a realização de análises *ad hoc*, através das ferramentas e dados armazenados. Esta é uma necessidade comum à maioria dos sistemas de

vigilância epidemiológica, que devem ser dotados de capacidade técnicas e de um repositório de dados que permitam responder às solicitações de outros estudos relacionados com esta temática ou de situações excepcionais fora do âmbito da vigilância epidemiológica automática de síndromes. Várias destas análises foram já desenvolvidas durante a implementação do SESP, algumas das quais serão também apresentadas no capítulo quinto.

No decorrer deste processo de definição do problema e objectivos, elaborou-se também um modelo conceptual do sistema, obtendo-se uma definição muito inicial mas robusta da arquitectura que se pretende para suportar tecnicamente a produção e disponibilização dos seus produtos (figura 12).

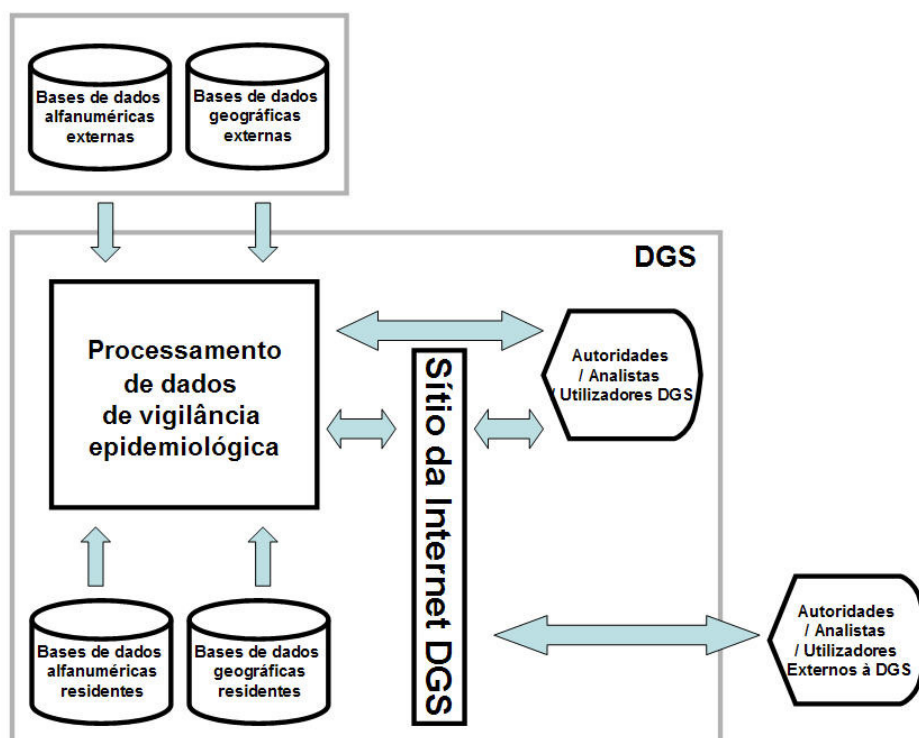


Figura 12 – Modelo Conceptual do SESP.

Estes objectivos constituíram também as especificações funcionais do sistema, e possibilitaram, a partir deste ponto, iniciar o processo de concepção técnica necessário para a sua implementação física.

4.4 Especificações técnicas para a implementação do sistema

Após a definição das especificações funcionais do sistema, o passo seguinte consistiu em delinear separadamente os vários componentes previstos, nomeadamente as bases de dados alfanumérica (relacional) e geográfica, bem como a arquitectura que as permite integrar.

4.4.1 Modelação da base de dados relacional

Uma definição abrangente para a base de dados pretendida para o SESP pode ser obtida em Dawes *et al* (2005), que dividem as bases de dados em dois tipos essenciais:

- Processamento de transacções (*Transaccional Processing*). Ocorrem diariamente múltiplas transacções, as quais são em geral de curta duração, envolvendo cada uma, um volume reduzido de dados. Os dados armazenados sofrem mudanças constantes;
- Armazém de dados (*Data Warehouse*). Ao invés do tipo anterior, as transacções acontecem em número reduzido, e envolvem grande volume de dados de cada vez. A maior parte das operações que incidem sobre estes dados são de consulta, e não de manipulação.

A base de dados aqui apresentada corresponde a um desenvolvimento do tipo *Data Warehouse*, uma vez que a recepção de dados envolve um grande volume, ocorre uma vez por dia, tal como o seu processamento, e os utilizadores finais efectuem essencialmente operações de consulta sobre a informação produzida.

Para delinear uma estrutura de dados capaz de dar resposta às necessidades expostas no capítulo anterior, recorreu-se ao método de modelação *ER (Entidade - Relacionamento)* para conceber, numa primeira fase, a estrutura do modelo de dados a adoptar. Este método consiste num modelo de dados puramente conceptual, destinado a representar o mundo real sob a forma de entidades e das associações que existem entre elas. A representação visual destes conceitos do modelo é efectuada recorrendo a um tipo específico de diagrama, designado *Diagrama de Entidade – Relacionamento*. As suas construções básicas são as *entidades*, *relacionamentos* e *atributos*. As *entidades* são conceitos, reais ou

abstractos, acerca da informação recolhida, os *relacionamentos* consistem em associações entre entidades e os *atributos* correspondem às propriedades que descrevem as entidades. A figura 13 apresenta o modelo ER delineado para o caso em análise. Os atributos das entidades são apresentados de seguida, existindo uma coincidência quase total entre os atributos das entidades e os incluídos nas tabelas derivadas deste modelo. A partir do modelo ER delineado, foram derivadas as relações (tabelas) constituintes da base de dados, de forma a iniciar a passagem de uma estrutura puramente conceptual para outra (modelo lógico) que precede a implementação física.

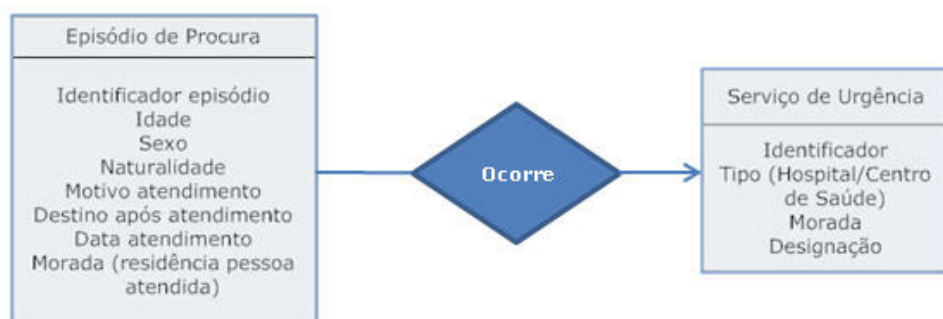


Figura 13 – Modelo Entidade – Relacionamento.

Por questões de economia de espaço, são apresentadas apenas as tabelas que constituem o “núcleo duro” da base de dados, existindo nesta muitas mais tabelas, associadas aos procedimentos automáticos diários de processamento de dados (questão clarificada mais à frente). A tabela 1 contém as tabelas da base de dados, sendo as respectivas características técnicas discriminadas no anexo 1. Fez-se também, em *Microsoft Access 2003*, um ensaio preliminar do esquema relacional (figura 14) da base de dados prevista.

Os dados alfanuméricos que alimentam actualmente o SESP, permitindo povoar estas tabelas, provêm de duas fontes: o INE (Instituto Nacional de Estatística), que contribui com dados sobre a população residente; os registos de procura de serviços de urgência, armazenados diariamente nos servidores de hospitais (sistema SONHO) ou centros de saúde (sistema SINUS). Esta componente “herdada” influenciou a estruturação da base de dados, pelo facto de existir informação produzida “a montante” e diariamente transferida para esta, cuja

estrutura relacional (das bases de dados implementadas nos servidores locais incluídos nos sistemas SONHO e SINUS) dificilmente pode ser contornada.

Tabela	Conteúdo
EURO_URGENCIAS_H	Afluência a urgências hospitalares – cada registo armazena uma consulta de urgência hospitalar.
EURO_URGENCIAS_H1	Afluência a urgências hospitalares – cada registo armazena uma consulta de urgência hospitalar, com acréscimo de uma coluna contendo o cálculo da idade da pessoa atendida
EURO_URGENCIAS_C	Afluência a urgências de centros de saúde – cada registo armazena uma consulta de urgência em centro de saúde
EURO_URGENCIAS_C1	Afluência a urgências de centros de saúde – cada registo armazena uma consulta de urgência em centro de saúde, com acréscimo de uma coluna contendo o cálculo da idade da pessoa atendida
EURO_INTERNAMENTOS	Episódios de internamento hospitalar - cada registo corresponde a um episódio de internamento hospitalar. Dados não utilizados presentemente.
EURO_INSTITUICOES	Características dos hospitais e centros de saúde incluídos no sistema
EURO_RESIDENCIAS	Definição do local de residência e naturalidade das pessoas atendidas
EURO_DESTINOS_C	Tipo de encaminhamento dado à pessoa após ida à urgência de centro de saúde
EURO_DESTINOS_H	Tipo de encaminhamento após ida à urgência hospitalar ou internamento
EURO_MOTIVOS_C	Motivo da ida a urgência de centros de saúde
EURO_MOTIVOS_H	Motivo da ida a urgência hospitalar
EURO_SEXO	Sexo da pessoa que ocorre às urgências

Tabela 1 – Tabelas da base de dados.

De facto, considerando como *normalização* o processo de estruturação dos dados no espaço do problema, que tem como objectivos essenciais eliminar a redundância e assegurar a flexibilidade (Riordan, 2005), esta situação teve como consequência a opção por não normalizar totalmente o modelo de dados, existindo ainda, por exemplo, tabelas redundantes no modelo (e.g. EURO_URGENCIAS_H e EURO_URGENCIAS_C poderiam ser juntas numa só tabela).

Caso se optasse por uma estrutura relacional inteiramente normalizada, as suas tabelas teriam de ser reconstruídas diariamente através de tarefas suplementares de processamento automático na base de dados residente do SSES, uma vez que as bases de dados que fornecem “a montante” os dados não correspondem ao modelo desejado. Esta situação obrigaria a uma sobrecarga adicional do esforço diário de processamento, levando, acima de tudo, ao aumento do tempo necessário para a execução das tarefas automáticas e atrasando o horário de disponibilização diária dos produtos. A secção 4.4.3 refere um pouco mais sobre esta situação.

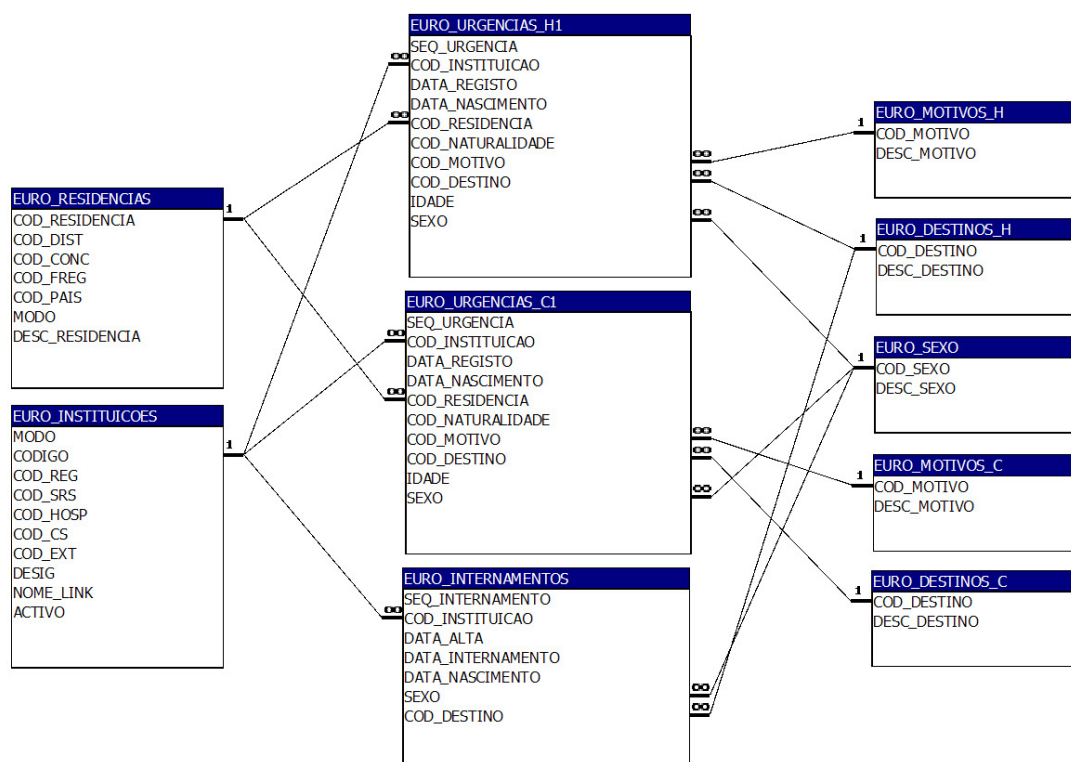


Figura 14 – Esquema relacional da base de dados do SESP.

4.4.2 Aquisição da base de dados geográfica

Para permitir a geração de *outputs* geográficos, nomeadamente mapas temáticos, e dotar o sistema da capacidade de realizar operações de análise espacial em geral, foram adquiridos vários conjuntos de dados geográficos (CDG) de diversos tipos. Nesta secção descrevem-se estes conjuntos de dados, colectivamente referidos como a base de dados geográfica do SESP, fornecendo metadados de referência, que podem ser consultados na tabela 2. Os dados geográficos de base provêm de sinergias entre institutos públicos, nomeadamente a DGS, o IGP (Instituto Geográfico Português) e o Instituto Nacional de Estatística (INE). O IGP fornece a Carta Administrativa Oficial de Portugal (CAOP), em formato vectorial, que define os limites administrativos de Portugal (distritos, conselhos e freguesias) e uma cobertura integral do país em ortofoto-imagens digitais com 50 cm de resolução espacial. Quanto ao INE, fornece a Base Geográfica de Referência de Informação (BGRI), da qual são utilizados vários CDG, destacando-se as cartografias de secções e subsecções estatísticas e de eixos de via à escala 1: 25 000 de todo o país, que complementam a informação anterior. Em termos formais,

esta informação geográfica corresponde ao conceito de *Foundation Databases* (Cromley e McLafferty, 2002), o qual se refere a CDG de base, que não foram produzidos especificamente por/para entidades relacionadas com Saúde Pública, podendo ser utilizados em múltiplos contextos.

Designação	Carta Administrativa Oficial de Portugal (CAOP)	Série Ortofotocartográfica Nacional	Base Geográfica de Referenciação Espacial (BGRI)
Escala de referência / resolução espacial	1/25 000	50 cm	1/25000
Data de referência	2008	2005	2001
Tema	Limites administrativos	-	Eixos Viários, Subsecções e Secções Estatísticas
Modelo de dados	Vectorial	Raster (imagem)	Vectorial
Representação espacial	Polígonos	Pixels	Linhas; Polígonos
Suporte	Shapefile	GeoTiff	Shapefile
Produtor	Instituto Geográfico Português	Instituto Geográfico Português	Instituto Nacional de Estatística
Sistema de referência	Datum 73,	Datum 73,	Datum Lisboa,
	Projecção Transversa de Mercator,	Projecção Transversa de Mercator	Projecção Transversa de Mercator,
	Origem das coordenadas no ponto central	Origem das coordenadas no ponto central	Origem das coordenadas no ponto fictício

Tabela 2 – Metadados da informação geográfica
(Carreira *et al*, 2007).

Por outro lado, a base de dados geográfica disponível conta ainda com as localizações pontuais de infra-estruturas de Saúde Pública (hospitais e centros de saúde), produzida pela DGS, e armazenada em suporte *shapefile*.

4.4.3 Arquitectura geral do sistema

Pretende-se que o SESP constitua um ambiente multiplataforma, integrando vários componentes tecnicamente muito diversos (figura 15). O ponto de entrada de dados no sistema ocorre no servidor de recolha, que acede diariamente e de forma automática a hospitais e centros de saúde para recolher os valores de afluência às urgências para uma base de dados *Oracle 9i* nele instalada.

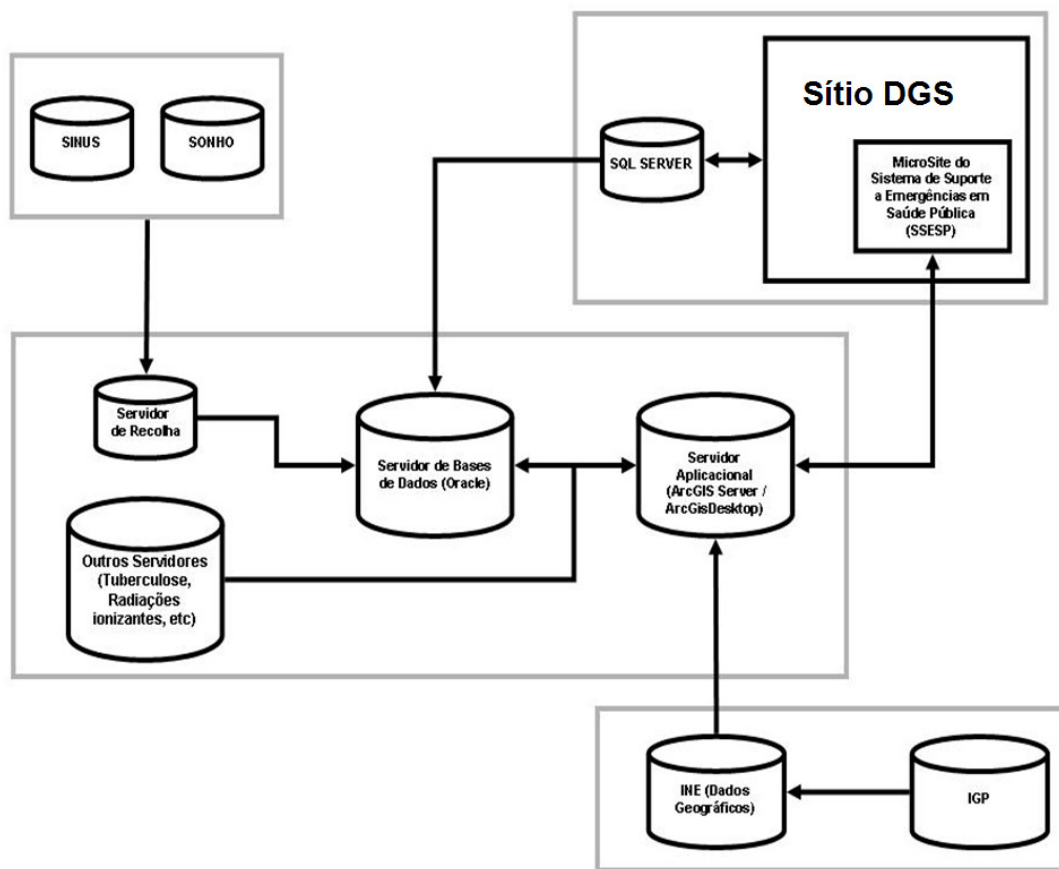


Figura 15 – Arquitectura do SESP
(Carreira *et al*, 2007).

Em termos técnicos, este é um servidor do tipo *Sun SPARC II*, dotado do sistema operativo *Sun Solaris 8*, tendo todas as rotinas que acedem aos hospitais e centros de saúde dos sistemas SONHO e SINUS sido programadas com recurso a ferramentas nativas deste sistema operativo. É aqui que reside realmente a componente “herdada”, atrás referida, uma vez que estes dados são recolhidos de acordo com uma estrutura relacional muito anterior à instalação do SESP, delineada de forma a permitir compatibilizar a comunicação e importação de dados a partir de bases de dados instaladas localmente e suportadas por SGBD muito diversos (na sua maioria *Oracle*, cujas versões variam entre *Oracle 6* e *Oracle 10g*). O cerne da arquitectura do SESP reside em dois servidores, idênticos em termos de *hardware* e sistema operativo (Processador *Intel Dual Xeon*, 10 *Gigabytes* de memória *RAM*, disco rígido com capacidade de armazenamento de 300 *Gigabytes*, Sistema Operativo *Microsoft Windows 2003 Server R2 32 bits*). Um destes

desempenha as funções de servidor de bases de dados, albergando uma base de dados *Oracle 10g*, a qual é responsável pela maior parte do processamento exercido sobre os diversos tipos de dados alfanuméricos, sendo que, após a recolha diária, os dados passam automaticamente do servidor de recolha para aqui. O segundo servidor, designado aplicativo, inclui todas as aplicações responsáveis pela geração dos produtos finais do sistema, desempenhando também funções de servidor de Internet, na medida em que suporta toda a infraestrutura do sítio na Internet dedicado ao SESP. Nele se encontram instaladas as seguintes aplicações de *software*:

- *ESRI ArcGIS Server 9.2*: sustenta a componente *WebGIS* do SESP, ou seja, a colocação de conteúdos geográficos na *Internet*;
- *ESRI ArcGIS 9.2*: permite suportar um sistema de informação geográfica do tipo mais comum, destinando-se à realização de análises “de gabinete”, bem como à preparação da informação que é disponibilizada no ambiente *WebGIS*;
- *Microsoft Visual Studio Professional 2005*: ambiente no qual foi efectuada a programação e o lançamento (*deployment*) do sítio na Internet dedicado;
- *Microsoft Internet Information Services (IIS) 6.0*: aplicação do tipo servidor de Internet (*Webserver*) que sustenta a disponibilização do sítio do SESP na Internet.

4.4.4 Conclusões

Esta fase do projecto revelou-se essencial e incontornável para a prossecução de vários objectivos. Entre estes, contam-se: o diagnóstico do estado da arte vigente na DGS em termos de actividades e sistemas de vigilância epidemiológica; as carências a ser colmatadas; a concepção geral do sistema e das respectivas características técnicas desejadas; o planeamento, quer das tarefas necessárias para colocar toda a arquitectura em funcionamento, quer da sequência de execução das mesmas. A parte do projecto a que se refere este capítulo foi talvez uma das mais diversificadas, na medida em que muitos interlocutores foram consultados, muitas soluções possíveis foram avaliadas (e por vezes previamente ensaiadas) e uma grande diversidade de rumos se apresentava. Foi ainda uma altura de muitas

decisões, por vezes difíceis, na medida em que acarretavam por vezes implicações de alguma importância para o futuro do SESP. A conclusão desta fase tornou possível avançar com alguma confiança para a implementação prática.

5 IMPLEMENTAÇÃO DO SISTEMA DE SUPORTE A EMERGÊNCIAS EM SAÚDE PÚBLICA

5.1 Introdução

Refere-se neste capítulo a implementação física dos componentes de vigilância epidemiológica diária do SESP. A segunda secção inclui a implementação da base de dados relacional que suporta toda a recepção e armazenamento de dados, sendo a terceira dedicada à implementação das tarefas diárias de processamento de dados. A quarta secção refere o *WebGIS* que ilustra geograficamente a informação obtida e a congregação de todos os elementos produzidos diariamente num sítio na Internet dedicado. A quinta secção refere estudos realizados numa óptica não diária, mas que fizeram uso da infra-estrutura montada para o SESP, e a última retira conclusões e indica funcionalidades cuja implementação se encontra prevista. Este tema foi também objecto da publicação de um *artigo* no 1º Congresso de Saúde e Tecnologias de Informação Geográfica (Carreira *et al*, 2007).

5.2 Implementação do repositório de dados alfanuméricos em Oracle 10g

Todas as tarefas de implementação e gestão do modelo de dados relacional em ambiente *Oracle 10g* (10.2.0.1) foram levadas a cabo recorrendo a duas ferramentas que constituem as alternativas mais comuns. Por um lado, recorreu-se a uma consola de administração proprietária, designada *Oracle Enterprise Manager 10g Database Control*, que centraliza e permite executar uma grande variedade de operações de gestão em modo gráfico. Por outro lado, utilizou-se *Oracle SQL Developer 1.2.0*, que consiste numa consola, instalada exteriormente à base de dados, para realizar, sobre esta, operações em modo de linha de comando, desde consultas variadas, passando pela geração de segmentos (designação da *Oracle* para os objectos da base de dados, tais como tabelas e vistas, entre outros) até operações de manipulação de dados, como inserção de registos. Esta aplicação suportou também todas as tarefas de desenvolvimento e ensaio dos blocos de código utilizados na automatização do processamento diário de dados, referidas na secção 5.3. Como já indicado, originalmente a maior parte das tabelas de base encontrava-se provisoriamente armazenada noutro servidor, em ambiente *Oracle 9i*

(9.2.0.1.0), desactivado no final da geração do SSESP. Começou-se, assim, pela análise da estrutura relacional da base de dados pré-existente, no sentido de planejar a migração de parte dos seus segmentos e procedimentos para o novo ambiente, mas também entender como estes iriam interagir com os novos componentes a lançar em *Oracle 10g*. Esta análise implicou aceder localmente ao servidor *Oracle 9i* com privilégios de administração (utilizador do tipo super administrador, ou *SYS*), para observar todo o conteúdo e estrutura da base de dados e realizar consultas ao dicionário de dados sempre que necessário. Este dicionário (aqui citado em sentido lato, pois a *Oracle* atribui-lhe uma definição mais restrita) consiste num conjunto de tabelas, pertencentes ao utilizador *SYS*, que contêm toda a informação de carácter interno acerca da base de dados, representando, no fundo, o repositório de metadados desta. Estas tabelas são apenas de consulta, podendo ser acedidas através de dois conjuntos de vistas predefinidas, *Data Dictionary Views* e *Dynamic Performance Views* (Dawes et al, 2005), na medida em que a estrutura real destas tabelas (nomes das tabelas e respectivos campos) se encontra codificada de forma muito pouco amigável para o utilizador. Estas consultas foram realizadas escrevendo directamente as interrogações em *Structured Query Language (SQL)*, através de uma consola de *SQL* nativa da própria base de dados, denominada *Oracle SQL Plus Worksheet*, e também em modo gráfico, através do ambiente de administração nativo de *Oracle 9i*, denominado *Oracle Enterprise Manager Console*.

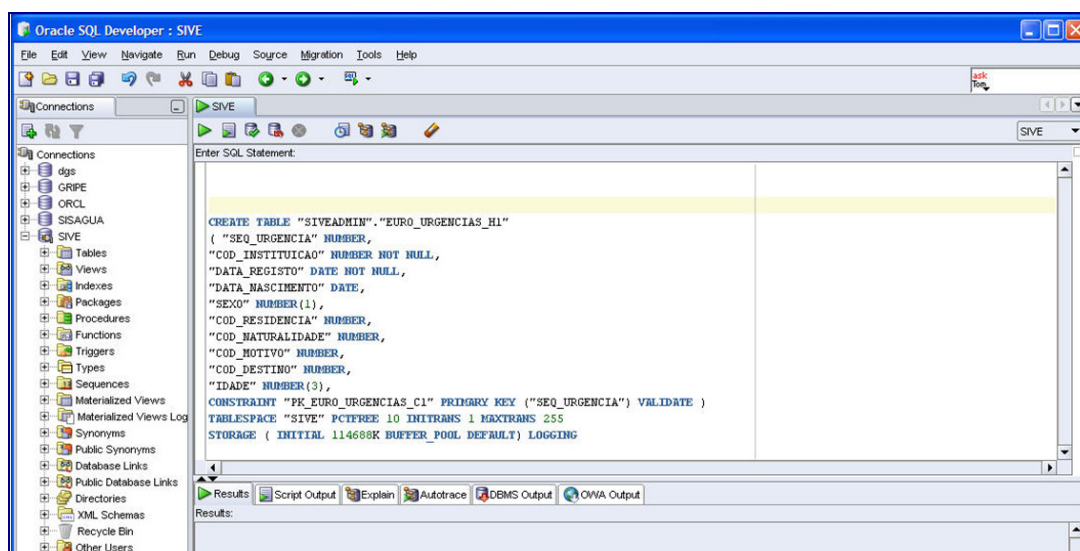


Figura 16 – Geração de tabela em *Oracle SQL Developer*.

Após a conclusão deste processo de análise, a primeira tarefa, já em *Oracle 10g*, consistiu em gerar as tabelas da base de dados, operação exemplificada (figura 16) em ambiente *Oracle SQL Developer* para a tabela EURO_URGÊNCIAS_H1.

Depois de criadas, as tabelas foram povoadas através de um *database link*. Este artifício permite ligar duas bases de dados de forma transparente, sendo possível invocar uma base de dados a que se pretende aceder a partir de outra, pela inclusão, no bloco de código SQL, de uma designação simples, lançada durante a geração do *database link*. A partir daqui, a introdução de dados nas tabelas, por vezes referida como *povoamento* (Dawes *et al*, 2005) foi efectuada em SQL com o comando *INSERT*, importando-se o conteúdo de cada tabela através da invocação do caminho para a base de dados de origem pela designação do *database link*, tendo sido escritos vários blocos de código SQL para o efeito, tais como:

```
INSERT INTO
SIVEADMIN.EURO_INSTITUICOES
SELECT * FROM
IGIF.EURO_INSTITUICOES @SISAGUA
```

Uma vez povoadas as tabelas, foram implementados três *packages*, receptáculos lógicos que permitem juntar blocos de código *PL/SQL*¹⁵ (*procedures*¹⁶ e/ou *functions*¹⁷), relacionados com uma mesma tarefa. Estes efectuem a transferência diária de dados do servidor de recolha para três das tabelas previamente geradas, que contêm todas as ocorrências registadas desde o início da recolha, no ano de 2003:

- EURO_URGENCIAS_H (procura de urgências hospitalares).
- EURO_URGENCIAS_C (procura de urgências em centros de saúde).
- EURO_INTERNAMENTOS (episódios de internamento hospitalar).

¹⁵ *Oracle Procedural Language for SQL*, linguagem de programação de procedimentos proprietária da *Oracle*. Dawes *et al* (2005).

¹⁶ Bloco de código *PL/SQL* guardado na base de dados que executa uma determinada acção. Dawes *et al* (2005).

¹⁷ Bloco de código *PL/SQL* guardado na base de dados que devolve um determinado valor quando é invocado numa instrução de *SQL*. Dawes *et al* (2005).

Estas tabelas, por conterem os maiores volumes de dados, são necessariamente as mais morosas de consultar. Assim, foram indexadas logo após o seu povoamento, em *Oracle Enterprise Manager 10g Database Control*, pelos campos mais frequentemente consultados, no sentido de melhorar o desempenho na velocidade de execução das consultas; a figura 17 indica, para a tabela EURO_URGENCIAS_H, os índices gerados e os campos sobre os quais incide a indexação.

Indexes		
		Number of Indexes: 4
Schema.Index	Indexed Columns	Column Position
SIVEADMIN.INDICE_URGENCIAS_H_1	COD_INSTITUICAO	1
	DATA_REGISTO	2
	COD_DESTINO	3
	SEXO	4
SIVEADMIN.INDICE_URGENCIAS_H_2	COD_INSTITUICAO	1
	DATA_REGISTO	2
	COD_DESTINO	3
	DATA_NASCIMENTO	4
SIVEADMIN.INDICE_URGENCIAS_H_3	COD_INSTITUICAO	1
	DATA_REGISTO	2
	COD_DESTINO	3
	COD_MOTIVO	4
SIVEADMIN.PK_EURO_URGENCIAS_H	SEQ_URGENCIA	1

Figura 17 – Indexação da tabela EURO_URGENCIAS_H.

Após serem povoadas nesta fase inicial, a maior parte das restantes tabelas de base sofre apenas manutenção quando necessário (exemplo da tabela EURO_INSTITUICOES, actualizada sempre que ocorrem reconfigurações na rede hospitalar), não sendo actualizadas diariamente.

5.3 Automatização do processamento de dados

Após o lançamento da estrutura relacional da base de dados, o desafio seguinte foi dotá-la da capacidade de processar dados de vigilância epidemiológica, o que implicou programar a realização diária de um conjunto de processos automáticos que suportam a produção de indicadores referentes ao dia anterior. Assim, a partir das 24 horas de cada dia, os dados produzidos nas várias instituições em vigilância começam a ser recolhidos pelo servidor de recolha, terminando este processo com a transferência dos dados para o servidor de bases de dados por volta das 8.45 horas da manhã. Minutos depois, inicia-se o processamento, que culmina por volta das 10.00 horas da manhã, com a actualização das tabelas finais que acolhem os indicadores que alimentam os produtos finais do sistema. Para além dos processos

diários, existem ainda outros que são executados uma vez por semana apenas, e cujos indicadores se referem aos sete dias anteriores.

Para tornar este fluxo possível, foram geradas várias tarefas agendadas de processamento diário de dados, executadas sequencialmente, bem como uma série de novas tabelas que recolhem os resultados dos vários passos. A criação e activação das tarefas foi efectuada em *Oracle Enterprise Manager 10g Database Control* com recurso à ferramenta *Database Scheduler*. Esta funcionalidade permite guardar separadamente os blocos de código *PL/SQL (Programs)* que executam operações sobre as tabelas e agrupá-los em sequências de execução com regras (*Chains*), como exemplificado na figura 18. Esta é uma abordagem modular que permite obter uma maior flexibilidade na gestão das tarefas a executar na base de dados, pois os blocos de código são geridos separadamente, sendo possível, por exemplo, utilizá-los várias vezes em *Chains* diferentes, alterar a ordem de execução, etc.

Steps

Step Name	Type	Object Name	Skip	Pause
STEP01	PROGRAM	SIVEADMIN.FLUXO1_P1	FALSE	FALSE
STEP02	PROGRAM	SIVEADMIN.FLUXO1_P2	FALSE	FALSE
STEP03	PROGRAM	SIVEADMIN.FLUXO1_P3	FALSE	FALSE
STEP04	PROGRAM	SIVEADMIN.FLUXO1_P4	FALSE	FALSE
STEP05	PROGRAM	SIVEADMIN.FLUXO1_P5A	FALSE	FALSE
STEP06	PROGRAM	SIVEADMIN.FLUXO1_P5B	FALSE	FALSE
STEP07	PROGRAM	SIVEADMIN.FLUXO1_P5C	FALSE	FALSE
STEP08	PROGRAM	SIVEADMIN.FLUXO1_P5D	FALSE	FALSE
STEP09	PROGRAM	SIVEADMIN.FLUXO1_P5E	FALSE	FALSE
STEP10	PROGRAM	SIVEADMIN.FLUXO1_P5F	FALSE	FALSE

Rules

Name	Condition	Action	Description
SCHED_RULES29	STEP01 COMPLETED	START "STEP02"	
SCHED_RULES31	STEP02 COMPLETED	START "STEP03"	
SCHED_RULES33	STEP03 COMPLETED	START "STEP04"	
SCHED_RULES34	STEP04 COMPLETED	START "STEP05"	
SCHED_RULES35	STEP05 COMPLETED	START "STEP06"	
SCHED_RULES36	STEP06 COMPLETED	START "STEP07"	
SCHED_RULES82	STEP07 COMPLETED	START "STEP08"	
SCHED_RULES83	STEP08 COMPLETED	START "STEP09"	
SCHED_RULES84	STEP09 COMPLETED	START "STEP10"	
SCHED_RULES85	STEP10 COMPLETED	START "STEP11"	

Figura 18 – Programação de Chain em Oracle Enterprise Manager 10g Database Control.

O agendamento (hora de execução, intervalo de repetição, etc.) é programado sob a forma de uma tarefa (*Job*) que invoca directamente *Chains* previamente geradas,

agendando a sua execução. Para obter todos os produtos pretendidos, foram desenvolvidos 11 *Jobs*, contendo no total 762 blocos de código *PL/SQL*. Como exemplo, os blocos de código envolvidos em parte dos passos do *Job 1* podem ser consultados no anexo 2, por ordem de execução. A sequência de passos de cada *Job* foi modelada visualmente em *Mindjet Mind Manager 2002*, no sentido de facilitar a sua gestão, numa sequência do tipo *tabela* → *operação* → *tabela*, na qual é possível associar texto a cada objecto, ficando armazenadas para consulta as características das tabelas e o bloco de código associado a cada operação, o que facilita bastante as tarefas de gestão do fluxo de tarefas, de cada vez que é necessário gerar um novo produto do sistema, ou alterar de alguma forma os existentes. Um exemplo do ambiente *Mind Manager* pode ser observado na figura 19, que ilustra parte da sequência de execução do *Job 1*, e o bloco de código que é executado no decorrer do passo assinalado. O anexo 2 inclui também uma imagem do *Job 1* em toda a sua extensão.

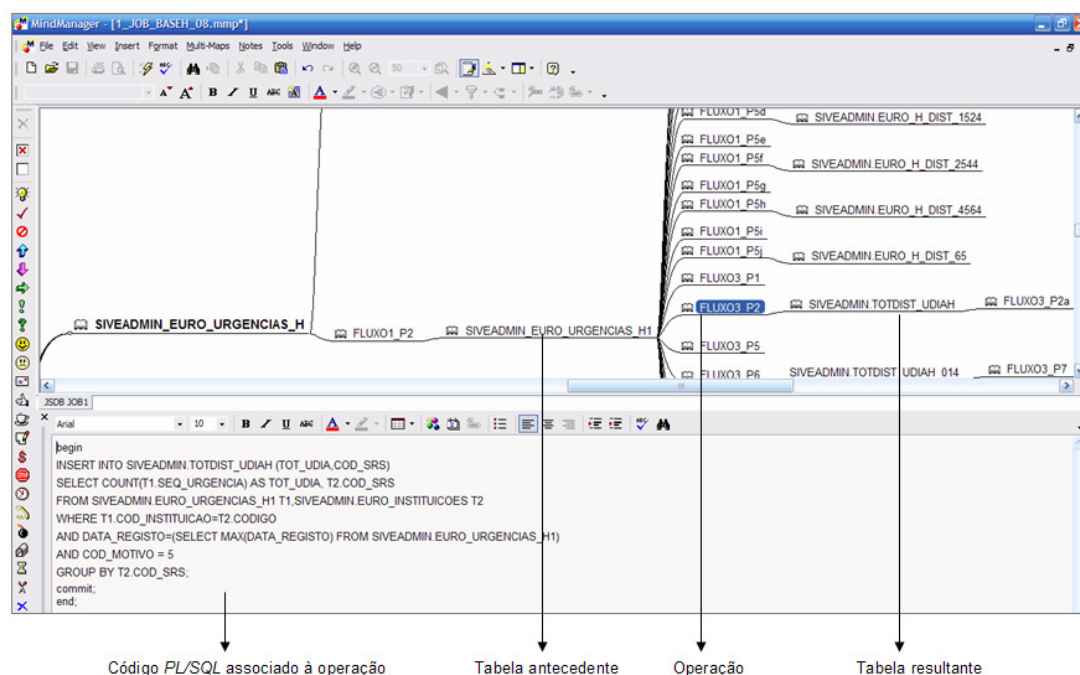


Figura 19 – Aspecto de modelação de um *Job* em ambiente *Mind Manager*.

Neste passo, as tabelas *EURO_URGENCIAS_H1* e *EURO_INSTITUICOES* (tabela não visível nesta imagem, mas no anexo 2) são interrogadas, e os dados resultantes são inseridos na tabela *TOT_DIST_UDIAH*, que surge de seguida na

sequência, contendo os totais de procura de urgências hospitalares por distrito no dia anterior, todas as idades.

Foi ainda necessário programar e implementar separadamente na base de dados, três procedimentos, que suportam a emissão diária de alertas epidemiológicos por correio electrónico, os quais são invocados diariamente pelo último *Job* a ser executado, após todas as tabelas de dados de actualização diária se encontrarem preenchidas.

5.4 Interfaces de consulta à base de dados

Uma vez estabelecida a produção diária de indicadores de vigilância epidemiológica, foram desenvolvidos os mecanismos que permitem aceder e consultar estes resultados, Para o efeito, foi gerado um sítio na Internet dedicado (figura 20), fisicamente alojado no servidor aplicacional, que centraliza a maior parte das consultas aos produtos finais do sistema, programado na linguagem C# em ambiente *Microsoft Visual Studio Professional 2005*. O acesso a este sítio é efectuado através de uma hiperligação localizada num sector específico do sítio na Internet da DGS, sujeito a autenticação por palavra-chave, e de acesso restrito aos decisores em Saúde Pública. As tabelas da base de dados que recolhem os valores actualizados dos indicadores encontram-se ligadas de forma permanente aos vários produtos finais do sistema (figura 21), através da instalação de um cliente *Oracle 10g* no servidor aplicacional, seguida de configuração de uma ligação ODBC.



Figura 20 – Página de entrada no sítio na Internet de acesso ao SSESP (DGS, 2008).

A actualização dos valores nestas tabelas reflecte-se de forma imediata nos produtos de análise e apoio à decisão. Após o acesso pela hiperligação *Urgências* (figura 20), esta página encontra-se dividida em duas partes, servindo a parte superior (figura 22) para a escolha do produto que se deseja visualizar e a parte inferior para apresentar o resultado.

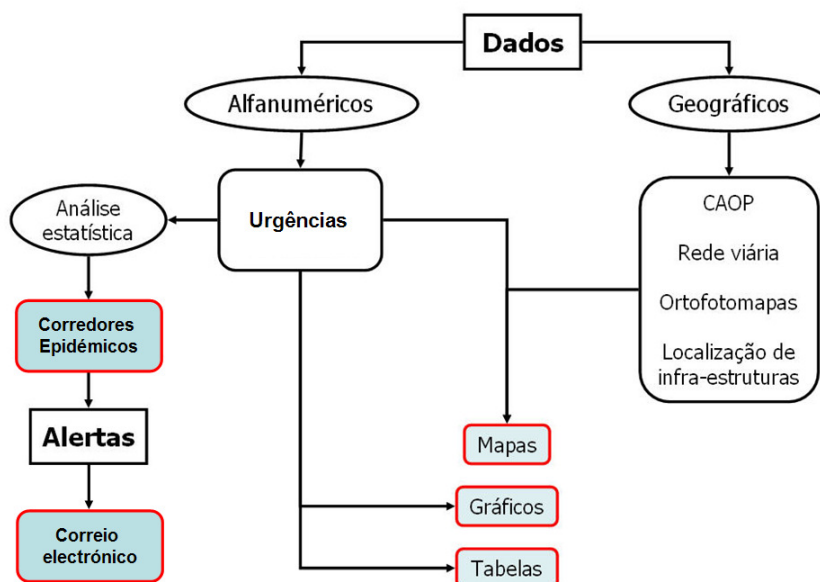


Figura 21 – Esquema geral dos produtos disponibilizados no âmbito do SESP.

Os produtos geográficos finais consistem em mapas coropletos (figura 23), sendo a representação geográfica feita por intervalos de intensidade da procura de cuidados de saúde em serviços de urgência por distrito no dia anterior e segundo uma escala de 5 cores graduadas.

Sistema de Suporte a Emergências em Saúde Pública

Procura de cuidados de saúde nos serviços de urgência, de acordo com os critérios:

Instituições	Unidade territorial	Escalões etários	Visualização
<input type="radio"/> Centros de Saúde <input type="radio"/> Hospitais <input checked="" type="radio"/> Ambos	<input checked="" type="radio"/> Total nacional <input type="radio"/> Aveiro <input type="radio"/> Beja <input type="radio"/> Braga <input type="radio"/> Bragança <input type="radio"/> Castelo Branco <input type="radio"/> Coimbra <input type="radio"/> Évora <input type="radio"/> Faro <input type="radio"/> Guarda <input type="radio"/> Leiria <input type="radio"/> Lisboa <input type="radio"/> Portalegre <input type="radio"/> Porto <input type="radio"/> Santarém <input type="radio"/> Setúbal <input type="radio"/> Viana do Castelo <input type="radio"/> Vila Real <input type="radio"/> Viseu	<input checked="" type="radio"/> Todas as idades <input type="radio"/> 0 a 14 anos <input type="radio"/> 15 a 24 anos <input type="radio"/> 25 a 44 anos <input type="radio"/> 45 a 64 anos <input type="radio"/> Mais de 64 anos	<input checked="" type="radio"/> Gráfico <input type="radio"/> Tabela <input type="radio"/> Mapa

ver

Figura 22 – Página de visualização dos produtos do SESP (DGS, 2008).

Os valores apresentados consistem no rácio da procura pela população residente, (multiplicado por 100.000 habitantes), no sentido de evitar a aparente concentração

de ocorrências junto ao litoral, onde habita a maior parte da população, caso fossem cartografados apenas os valores em bruto. A análise por 100.000 habitantes é frequentemente citada em textos de Epidemiologia (Beaglehole *et al*, 1993), por facilitar a análise de resultados nas situações em que os totais de ocorrências de um determinado fenómeno apresentam valores muito reduzidos face ao total da população em estudo.

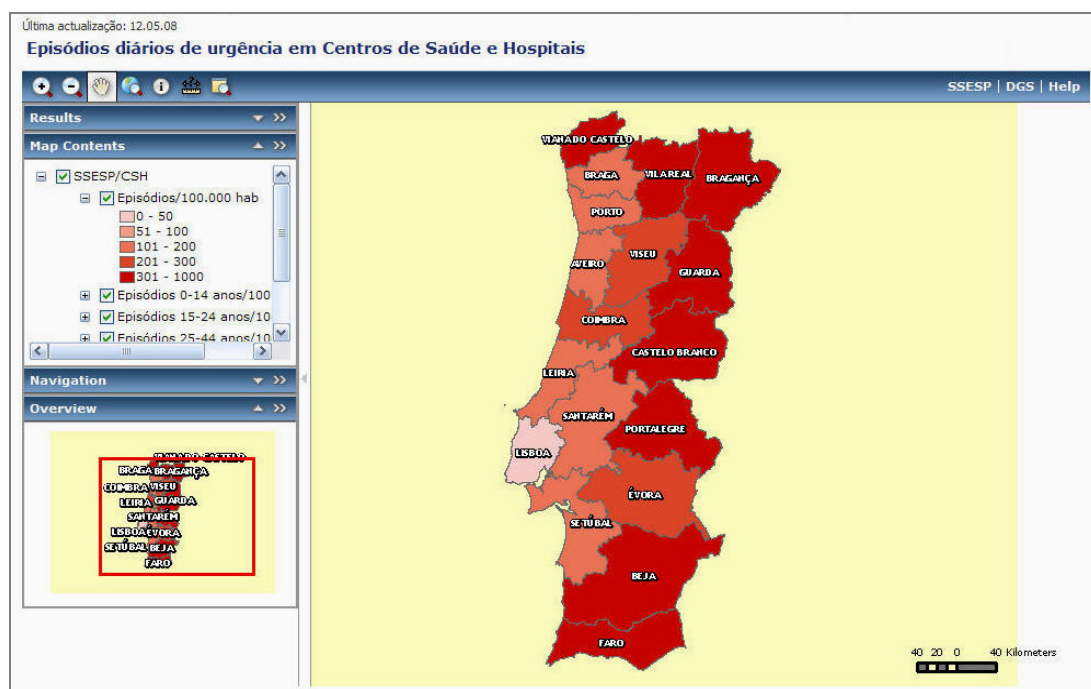


Figura 23 – Procura distrital de urgências por 100.000 habitantes em 12-05-2008 (DGS, 2008).

A preparação dos mapas temáticos iniciou-se com a invocação da ligação *ODBC* preestabelecida à base de dados em *ESRI ArcCatalog 9.2*, através da ferramenta *Database Connections*. Esta operação permite a utilização das tabelas da base de dados *Oracle* contendo os valores diários de procura em ambiente *ESRI ArcMap 9.2*, procedendo-se à ligação destas tabelas à cartografia de base, através de uma ligação relacional (*Join*). Já em ambiente *ESRI ArcGIS Server 9.2*, o ficheiro de projecto gerado no passo anterior para cada mapa temático (*.mxd*) foi invocado para a geração de um *serviço*, resultando desta operação uma página na Internet contendo o mapa desejado e algumas funcionalidades SIG de apoio à sua visualização (aumentar e diminuir a escala, entre outras), que pode ser acedida através de um qualquer *browser* (*Microsoft Internet Explorer*, por exemplo). Como parte do processo, este *serviço* foi, por sua vez, invocado em *Microsoft Visual*

Studio Professional 2005 para reconfigurar algumas características gráficas da página (recorrendo à linguagem C#), e integrá-la no sítio na Internet que alberga os produtos do sistema. Este tipo de análise permite obter uma visão diária da situação epidemiológica em Portugal Continental, fornecendo uma comparação célere entre unidades territoriais.

Todavia, Beaglehole *et al* (1993) refere que a análise epidemiológica incide sobre o tempo (no sentido da análise da evolução temporal dos fenómenos), o número de pessoas afectadas, o espaço (geográfico), e as classes de pessoas afectadas. A análise isolada dos acontecimentos diários revela-se assim insuficiente, sendo igualmente importante para a detecção e acompanhamento de surtos, a monitorização da evolução temporal da procura, assegurada no SSESF através de gráficos de linhas (figura 24), que apresentam a evolução distrital da procura diária no ano a decorrer e no ano anterior.

Para o desempenho de funções de vigilância epidemiológica, a detecção precoce de surtos é essencial para desencadear, adequada e atempadamente, acções de investigação e controlo. Neste contexto, uma questão fundamental prende-se com a escolha dos valores relativos ao número de casos, a partir dos quais se justifica desencadear alertas. Para tal, o SSESF baseia-se no método dos corredores endémicos de Bortmann (1999)¹⁸, o qual permite representar graficamente quatro zonas que, classificadas em função da gravidade da situação, se podem denominar: “Zona de sucesso”, “Zona de segurança”, “Zona de alerta” e “Zona Epidémica”. Sobre as zonas, é desenhada a linha correspondente ao número de casos semanais notificados durante o ano em estudo. Recorrendo a este método, o SSESF fornece dois dos produtos finais, cuja interpretação deve ser conjunta: um corredor epidémico (terminologia adoptada no âmbito do SSESF) nacional (figura 25), calculado semanalmente; a emissão automática de alertas por correio electrónico, avaliada diariamente.

¹⁸ Bortmann (1999) propõe um método de análise e representação gráfica com auxílio de uma vulgar folha de cálculo. Este método baseia-se na construção de “corredores endémicos”, calculados com base nas séries de valores semanais dos 5-7 anos anteriores ao ano em análise. Calcula-se uma média geométrica do número de casos por unidade de tempo e respectivos limites do intervalo de confiança (IC) de 95%. A representação gráfica das linhas correspondentes à variação temporal da média e dos limites do IC de 95% delimita 4 áreas, designadas por “corredores endémicos”, correspondendo a 4 zonas que, classificadas em função da gravidade da situação, se podem denominar: “Zona de sucesso”, “Zona de segurança”, “Zona de alerta” e “Zona Epidémica”. Sobre estas zonas, é desenhada a linha correspondente ao número de casos semanais notificados durante o ano em estudo. Segundo este autor, este modelo é útil para a detecção precoce de surtos de algumas doenças endémicas de evolução aguda e período de incubação curto, desde que o intervalo de tempo entre cada ocorrência e a sua notificação seja mínimo.

Evolução da procura diária de urgências em hospitais, todas as idades, ambos os sexos, no distrito de Beja

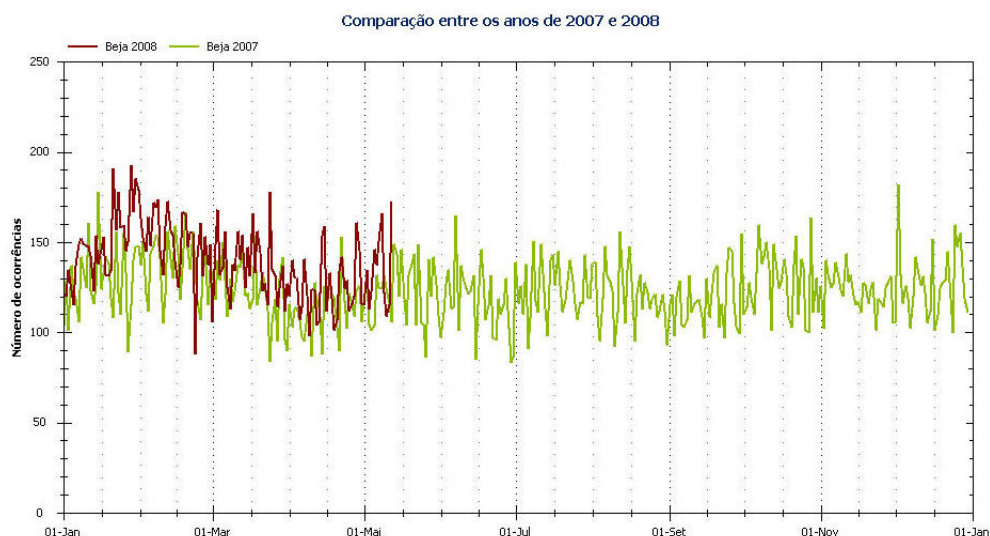


Figura 24 – Evolução da procura, Distrito de Beja, 2007 e 2008 (DGS, 2008).

Cada um destes produtos apoia-se na comparação dos totais nacionais de procura com valores limiares (obtidos a partir de dois corredores epidémicos previamente calculados: semanal e diário), efectuada com periodicidade semanal e diária.

Evolução da procura diária de urgências em centros de saúde e hospitais, todas as idades, ambos os sexos, em Portugal Continental

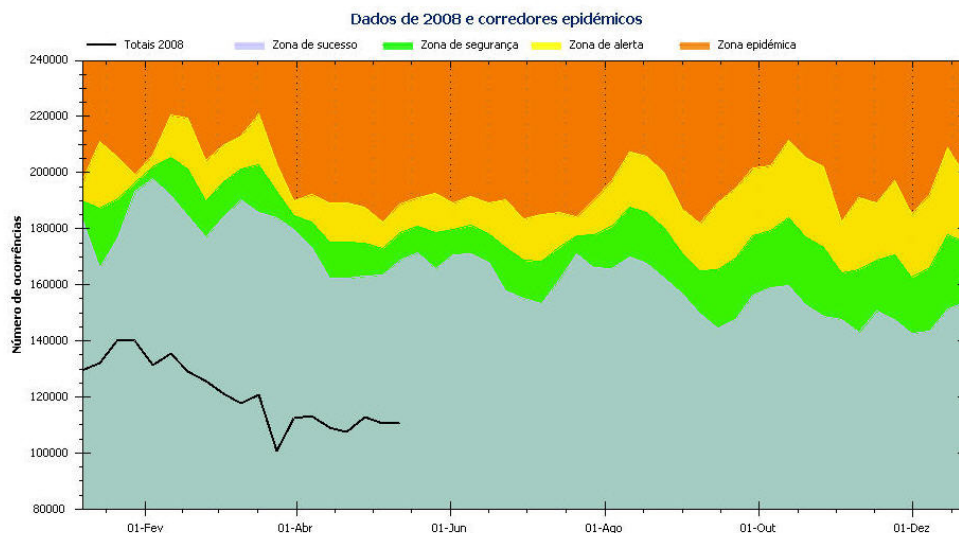


Figura 25 – Corredor epidémico nacional semanal, ano de 2008 (DGS, 2008).

Assim, a utilização conjunta destes dois produtos implica que a emissão de um alerta (efectuado sempre que, no corredor epidémico diário, seja ultrapassado o

limiar da Zona de Alerta) deva ser acompanhada pela observação da evolução da curva semanal contida na figura 25, na medida em que o aumento sustentado do seu declive pode indiciar uma epidemia.

Por último, são disponibilizados ainda quadros contendo os valores totais da procura diária no ano corrente e no ano anterior, para completar a tarefa de interpretação dos meios gráficos e geográficos de análise, quando se revelam insuficientes, e também para apoiar estudos estatísticos de fundo com base nestes dados, na medida em que os quadros podem ser copiados directamente para outro *software* (e.g. uma folha de cálculo) para tratamento posterior.

5.5 Outras aplicações do SESP

Para além da sua componente de vigilância epidemiológica diária, a infra-estrutura em que se apoia o SESP permite suportar a realização de outros estudos no âmbito da Geografia da Saúde, graças às ferramentas e dados armazenados de que dispõe. Estes estudos são realizados numa base pontual, e representam extensões da utilidade do sistema face ao que foi apresentado nas secções anteriores deste capítulo.

A secção seguinte inclui actividades de apoio à decisão realizadas internamente à DGS com o objectivo de reforçar a vigilância epidemiológica diária no decorrer de períodos epidémicos. As três secções restantes referem estudos de fundo em Geografia da Doença e Geografia dos Sistemas de Cuidados de Saúde, que tiraram partido da infra-estrutura montada.

5.5.1 Actividades de reforço do apoio à decisão em vigilância epidemiológica

Estas tarefas estão intimamente relacionadas com o apoio à decisão em vigilância epidemiológica, mas, por limitações técnicas e/ou determinações institucionais, não podem ser disponibilizadas remotamente. Um exemplo consiste na delimitação de um cordão sanitário com recurso à ferramenta *ESRI ArcGIS Network Analyst 9.2*, utilizando a localização de um foco epidémico e a rede viária da área, tendo como critério um determinado limiar, que pode ser uma distância percorrida por estrada ou um tempo de viagem (figura 26).

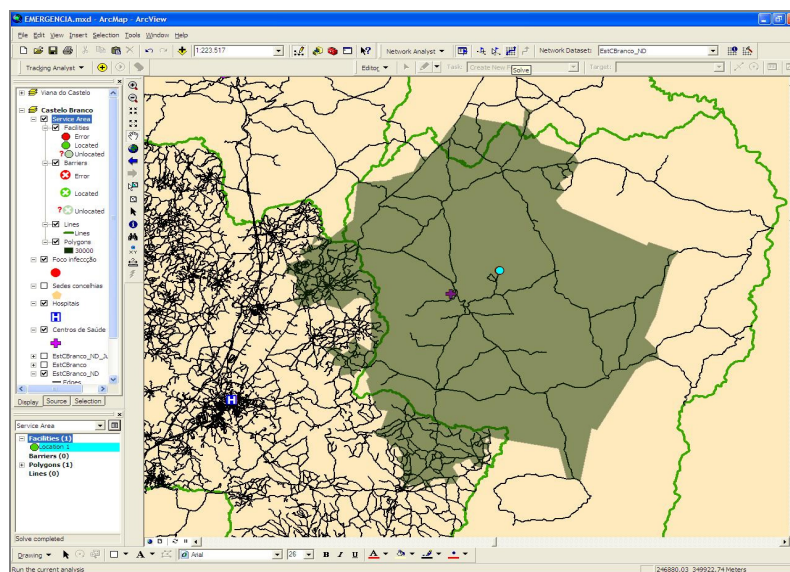


Figura 26 – Cordão sanitário gerado a partir da localização de um foco epidémico.

Outro tipo de utilização consiste na análise multitemporal da procura de serviços de urgência numa determinada área. Consiste, no essencial, na observação da distribuição espacial da procura na área de estudo, ao longo de um determinado intervalo temporal. Este estudo foi ensaiado na DGS com os dados da epidemia sazonal de gripe de 2007, utilizando-se para o efeito a ferramenta *ESRI ArcGIS Tracking Analyst 9.2*, que permitiu gerar o “filme” da evolução temporal da procura ilustrado na figura 27; valores elevados são representados a vermelho, os valores baixos a verde.

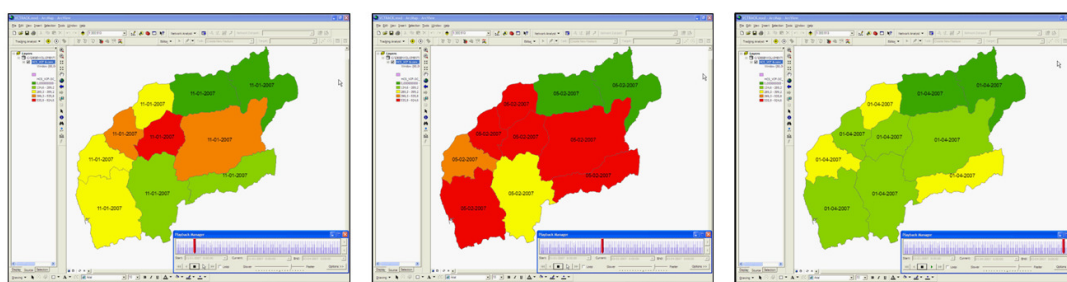


Figura 27 – Evolução da procura de serviços de urgência, dias 11-1-2007, 05-02-2007 e 01-04-2007, concelhos do distrito de Viana do Castelo.

Este tipo de análise permite determinar facilmente *quando* e *onde* ocorreu o pico da epidemia, e como evoluiu a procura destes serviços no território, facilitando a compreensão dos canais de comunicação da doença. Esta é uma análise

essencialmente *retrospectiva*, cuja vantagem reside no planeamento de intervenções e medidas futuras.

5.5.2 Amostragem espacial na avaliação do nível de saúde e acessibilidade aos cuidados de saúde nas comunidades imigrantes

Este estudo integra-se num projecto conjunto da Faculdade de Medicina da Universidade de Lisboa e DGS, com o objectivo de aferir o estado de saúde das comunidades imigrantes residentes em Portugal (de origem africana e brasileira) e respectivos níveis de acesso aos serviços de saúde, tendo sido também apresentado no 1º Congresso de Saúde e Tecnologias de Informação Geográfica (Carreira, 2007b).

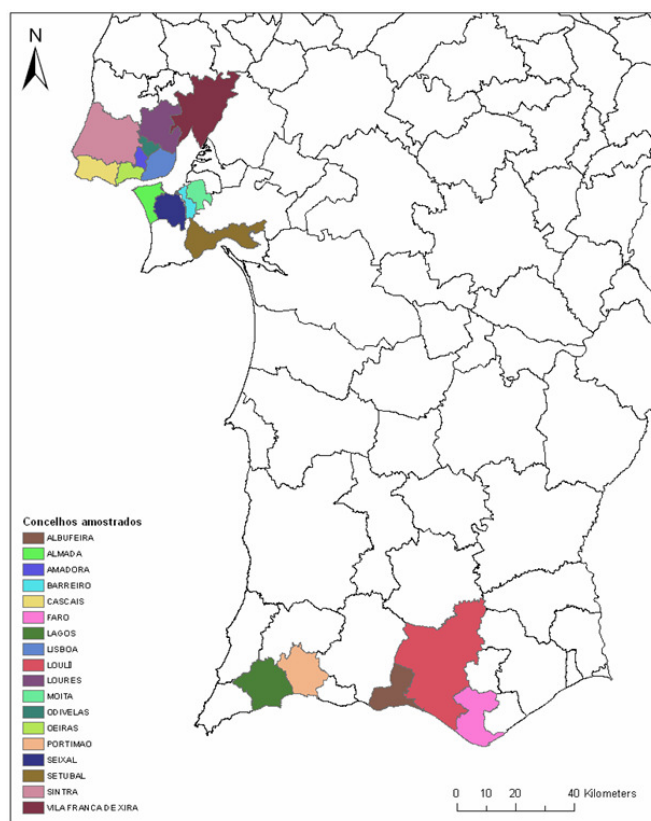


Figura 28 – Concelhos amostrados
(Carreira *et al*, 2007b).

A área de estudo correspondeu a 18 concelhos de Portugal Continental (figura 28), definidos, com base num estudo prévio, como sendo aqueles que concentravam a maioria dos indivíduos destas comunidades; segundo o mesmo estudo, foram

também escolhidas dentro destes concelhos apenas as freguesias mais representativas. Esta constitui uma investigação de carácter epidemiológico no âmbito da Geografia dos Sistemas de Cuidados de Saúde (Boulos *et al*, 2001), lidando com o planeamento e gestão dos serviços de saúde, após diagnosticar as necessidades da população-alvo em termos de cuidados de saúde.

A recolha de dados acerca da população-alvo é feita através de questionário, realizado ao domicílio por equipas de duas pessoas, originárias das comunidades em estudo. Todavia, a dimensão e distribuição geográfica da população-alvo são, no limite, desconhecidas, e mesmo que assim não fosse, a capacidade logística e financeira do projecto inviabilizaria a possibilidade de entrevistar a totalidade dos indivíduos pertencentes a estas comunidades. Tentando ultrapassar estas limitações, optou-se pela selecção aleatória das habitações a entrevistar. Para representar eficazmente a população em estudo, a amostra tem de ser significativa, o que depende essencialmente da dimensão amostral e do método de amostragem. Quanto à dimensão amostral, a sua escolha teve a ver com as possibilidades logísticas e financeiras do projecto. No que respeita ao método de amostragem, uma vez definida a área de estudo, o processo de recolha da amostra passou a ser orientado por um método de amostragem *espacial*, tomando como referência Stehman (1999), que discute a espacialização dos métodos de amostragem e suas características essenciais. Escolheu-se assim implementar uma *amostragem aleatória espacial por clusters*, num passo apenas, em que a selecção dos *clusters* obedece a uma *amostragem espacial aleatória simples*. O recurso a este método surge como forma de evitar, na medida do possível, uma das formas mais importantes de erro sistemático, o *viés de selecção*, o que aconteceria se a realização das entrevistas fosse confinada a áreas muito restritas, onde se soubesse de antemão que habitam as comunidades em estudo (e.g. certos bairros), correndo-se o risco da amostra ser pouco representativa da população-alvo.

5.5.2.1 Desenvolvimento da metodologia de amostragem

Várias operações de análise espacial foram utilizadas para operacionalizar a amostragem espacial pretendida, permitindo seleccionar as habitações a serem visitadas pelos entrevistadores e facilitar a orientação no terreno durante as respectivas visitas, no decurso do trabalho de campo. Foram utilizados três tipos de dados pertencentes à base de dados geográfica do SESP, nomeadamente os

limites administrativos da CAOP, a cobertura nacional ortofotocartográfica e os eixos viários da BGRI. As operações foram efectuadas em ambiente *ESRI ArcGIS (ArcView)* 9.2. A figura 29 apresenta o fluxo das operações efectuadas em cada concelho isoladamente, tendo sido escolhido o concelho de Faro para ilustrar o processo. Este iniciou-se com a definição do sistema de referência do ambiente de trabalho (*ESRI ArcMap*), adoptando-se o *Datum Lisboa, Projecção Transversa de Mercator, Origem das coordenadas no ponto fictício*.

Após estes preparativos, seleccionou-se o polígono da área do concelho no respectivo tema, centrando-o na janela do mapa (*dataframe*). Procedeu-se de seguida à criação automática de outro tema, contendo a quadrícula de amostragem, de forma a cobrir todo o polígono do território concelhio, constituindo os polígonos desta quadrícula a *unidade amostral primária*, com dimensão de $50 \times 50 \text{m}^2$, denominados *clusters*. No caso do concelho de Faro, todas as freguesias foram consideradas elegíveis para amostragem, bastando utilizar o limite concelhio previamente escolhido, sendo este sobreposto à quadrícula de amostragem para seleccionar apenas os *clusters* cujo centro se encontra contido nos limites do concelho, os quais são exportados para um novo tema. Sobre este incide um processo de *amostragem espacial aleatória simples*, utilizando uma ferramenta específica, com o objectivo de seleccionar 20 clusters, que indicam os locais de recolha dos dados no terreno.

Uma vez que a quadrícula abrange todo o território, e não apenas as áreas ocupadas por habitações, é necessário validar por interpretação visual (*fotointerpretação*) os 20 clusters seleccionados, recorrendo a ortofoto-imagens digitais (figura 30), sendo cada *cluster* aceite ou recusado, consoante contenha pelo menos uma habitação, ou nenhuma, respectivamente. Os processos de amostragem e validação são repetidos várias vezes até serem obtidos 20 clusters válidos; em cada sucessiva tentativa é admitido um número de *clusters* correspondente aos rejeitados na amostragem anterior, sendo cada tentativa efectuada sobre uma nova quadrícula, previamente gerada por exclusão de todos os *clusters* seleccionados, aceites ou não, na amostragem anterior.

Após obtenção dos 20 *clusters* finais, inicia-se então a preparação de um relatório destinado aos entrevistadores, contendo imagens dos locais a visitar no terreno (as imagens dos *clusters* válidos são recolhidas no decurso da validação pelas ortofoto-imagens digitais), a localização relativa correspondente (o nome do eixo viário que atravessa o *cluster* ou aquele mais próximo), quando esta é passível

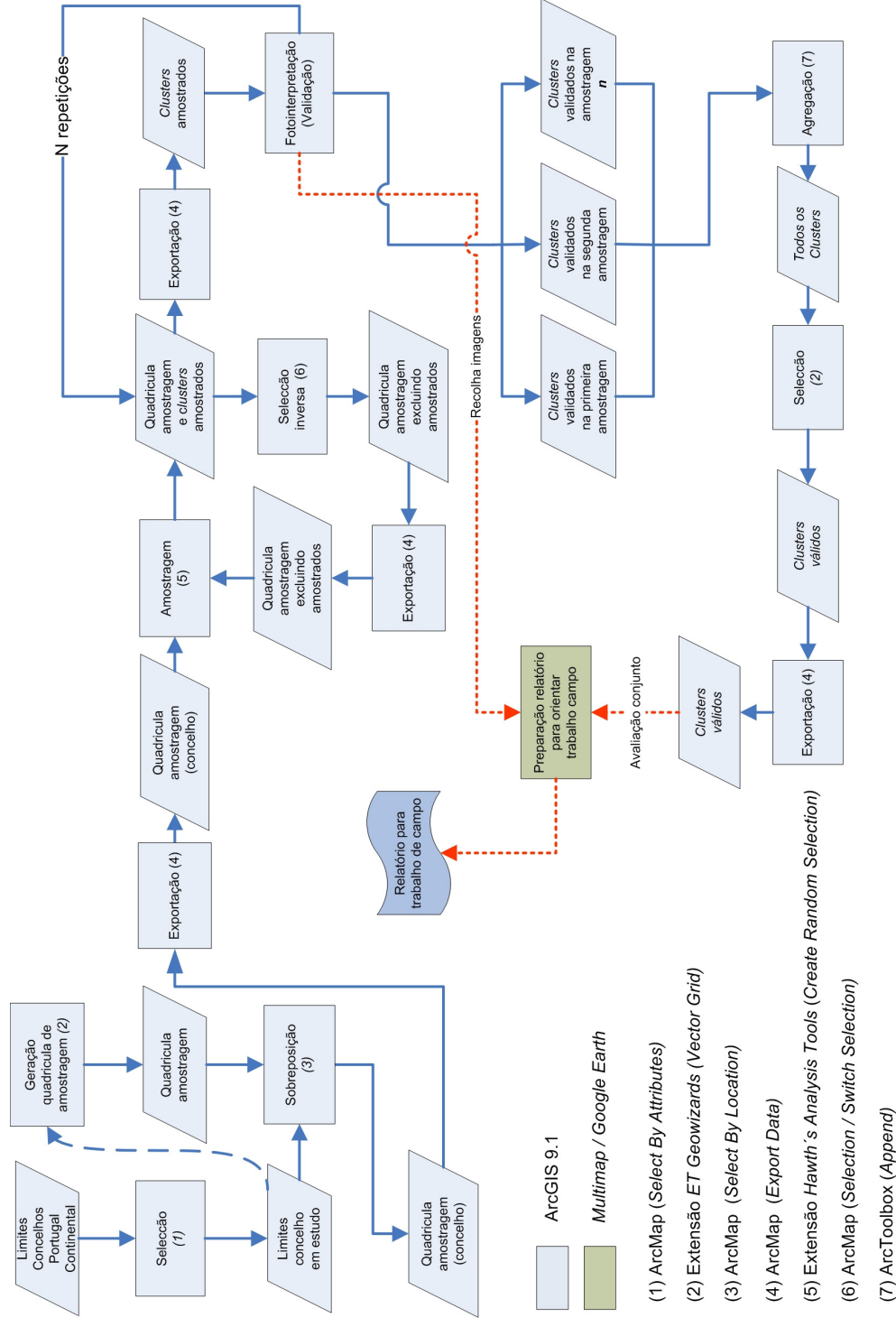


Figura 29 – Fluxograma das operações realizadas (Carreira et al, 2007b).

de ser obtida, e ainda uma proposta da ordem pela qual os locais podem ser visitados. As coordenadas dos centróides destes *clusters* são extraídas e carregadas em receptores *GPS*¹⁹, os quais permitem aos entrevistadores orientarem-se de forma eficaz no terreno, atingindo as localizações pretendidas com o mínimo dispêndio de tempo.

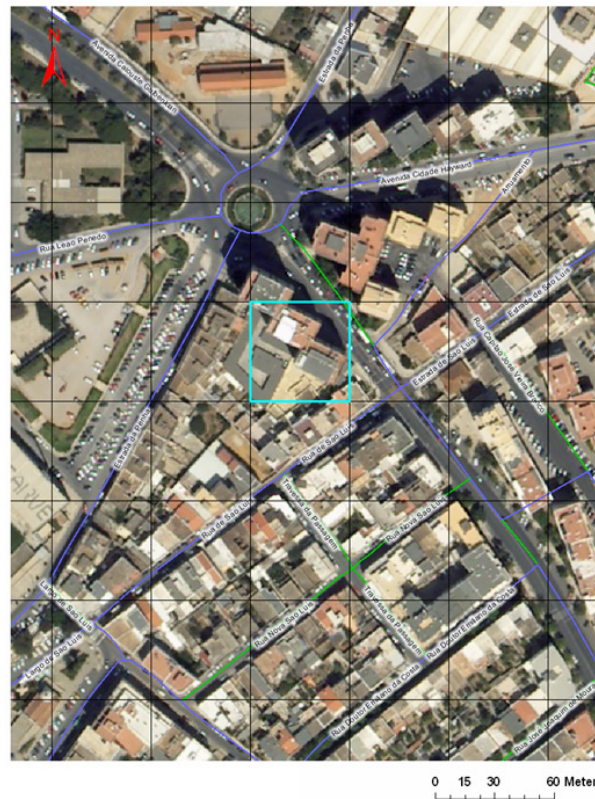


Figura 30 – Validação de um *cluster* da amostra (Carreira *et al*, 2007b).

Finalmente, o processo de colheita da amostra é realizado por equipas compostas por duas pessoas originárias das comunidades em estudo. As equipas de amostragem visitam então todos os domicílios (*unidade amostral secundária*) incluídos em cada *cluster*, convidando as pessoas elegíveis para a amostra a responder a um questionário em data posterior.

¹⁹ Sistema de Posicionamento Global (*Global Positioning System*). Corvallis (2000).

5.5.2.2 Resultados e considerações finais

Este é um estudo que ainda se encontra a decorrer, tendo sido abordados até ao presente 13 municípios, obtendo-se uma amostra com 3235 indivíduos, sendo 67% oriundos dos PALOP e 33% de origem brasileira. A taxa de adesão dos domicílios contactados foi de 97,9%.

As opções metodológicas tomadas oferecem ainda alguns pontos para discussão, no sentido do aperfeiçoamento da metodologia. A título de exemplo, o recurso a dados geográficos referentes à localização do edificado melhoraria a produtividade de todo o processo técnico de selecção dos clusters, tornando possível a exclusão automática das áreas desprovidas de habitações. Outras questões podem ainda ser consultadas em Carreira *et al* (2007b).

5.5.3 Ensaio de realocação de serviços de urgência com recurso a redes neuronais artificiais não supervisionadas

As recentes alterações da rede de prestação de serviços de urgência em Hospitais e Centros de Saúde de Portugal Continental originaram, entre outras reconfigurações, o encerramento de alguns serviços de urgência, o que levou à alteração da configuração territorial da oferta destes serviços. A situação levanta alguma curiosidade acerca da eficácia da actual rede de urgências face à configuração anterior, e da necessidade de optimização na localização destes recursos, assumindo que, por definição, são escassos. Surge assim a necessidade de encontrar métodos para avaliar o desempenho desta rede em diferentes momentos, mas também formas de modelar novas configurações territoriais potencialmente mais evoluídas.

Optimizar as localizações dos serviços pode ser entendido como uma tentativa de encontrar soluções para o problema “*p-median*”, o qual corresponde a um problema clássico de optimização combinatória. Este pode ser entendido como a localização de nós numa rede minimizando a soma de todas as distâncias de cada ponto dessa rede ao nó mais próximo (Henriques e Bação, 2004). Tenta-se aqui dar resposta ao problema recorrendo a uma Rede Neuronal Artificial (RNA), que consiste num conjunto de elementos de processamento simples (unidades ou neurónios) cuja funcionalidade é inspirada na estrutura neuronal do cérebro humano. Um tipo de RNA que tem sido usada com bastante êxito na análise de

dados espaciais são os *Self-Organizing Maps (SOM)*, que correspondem a redes neuronais artificiais capazes de organizar informação num espaço, normalmente bidimensional, usando um vector de n dimensões de neurónios (Kohonen, 2001).

O SOM é baseado em aprendizagem não supervisionada, o que significa que o treino é inteiramente orientado pelos dados e que os neurónios do mapa competem entre si. Escolheu-se usar o SOM neste trabalho, como uma aproximação à resolução do problema de encontrar novas configurações para a rede de urgências numa área de estudo em Portugal Continental, ou seja, tentar redefinir as localizações (nós) dos serviços de urgência de forma a posicionar de forma mais eficaz cada um destes serviços face à população residente na área de estudo. Por outro lado, uma vez determinadas localizações alternativas, e caso se revelem viáveis, outra possibilidade, embora não seja ainda desenvolvida neste estudo, consiste no (re)desenho das áreas de influência destes serviços com base nas novas localizações encontradas.

A secção seguinte apresenta uma breve revisão sobre o SOM, apresentando as suas principais características e o respectivo algoritmo. Em 5.5.3.2 faz-se uma breve introdução ao problema do *p-median* e desenho de regiões. As secções subsequentes apresentam a discussão metodológica deste estudo e discutem os resultados obtidos, respectivamente. A última secção inclui as notas finais e aponta perspectivas futuras para o desenvolvimento desta metodologia.

5.5.3.1 Self-Organizing Maps (SOM)

Como já referido, uma Rede Neuronal Artificial (RNA) consiste num conjunto de elementos de processamento simples, unidades ou nós, cuja funcionalidade é baseada numa definição grosseira da estrutura neuronal do cérebro humano. As RNA correspondem, assim, a programas de computador criados para simular o processo de conhecimento humano através do estabelecimento e reforço de ligações entre os dados de entrada e de saída (Henriques e Bação, 2004). Estes métodos consistem numa aproximação ao conceito de Inteligência Artificial, na medida em que tentam recriar algumas capacidades de processamento do cérebro humano, o que corresponde a tentar recriar a sua estrutura. Uma das características mais interessantes é a tentativa de modelação artificial das ligações entre os neurónios do cérebro humano (*sinapses*), a qual é efectuada através da

definição de uma força ou peso de ligação entre os nós da rede (os neurónios artificiais).

A utilidade dos SOM assume especial relevância nas tarefas de *clustering* e visualização. Um dos seus usos possíveis consiste em projectar grande quantidade de informação, de elevada dimensionalidade, num número menor de dimensões (normalmente uma ou duas), retendo todavia o padrão inicial da informação. Desta forma, informação que se encontra próxima de um conjunto de informação permanecerá próximo no *output*. Os neurónios são unidades sensíveis que possuem a capacidade de modificar um conjunto de parâmetros de forma a alterar o seu *output* durante o processo de aprendizagem (Henriques e Bação, 2004). No que respeita aos seus princípios básicos, o algoritmo de treino do SOM envolve essencialmente dois passos, o processo de aprendizagem da rede e a sua projecção.

A aprendizagem da rede diz respeito ao processo de movimentação dos neurónios, com o objectivo de criar um conjunto representativo dos dados de entrada. Esta fase reduz, normalmente, a complexidade nos dados, na medida em que o número de neurónios é inferior ao número de dados de entrada. Este pode, assim, ser entendido como o processo de classificação ou *clustering*.

O processo de projecção dos neurónios possibilita a representação numa rede regular, normalmente bidimensional, o que permite ao utilizador uma interpretação simplificada dos resultados (Figura 31). Esta fase é essencialmente de produção de uma forma visual de interpretação da posição dos neurónios permitindo assim ao utilizador a identificação dos *clusters* formados. A experiência do utilizador é o factor essencial nesta fase, estando a qualidade de interpretação da projecção muito dependente do seu discernimento.

O algoritmo SOM baseia-se nos passos anteriormente descritos, nomeadamente a aprendizagem e projecção dos neurónios. Para levar a cabo o treino de uma rede SOM é necessário especificar um elevado número de parâmetros, facto que decorre da natureza heurística do SOM e da ausência de um critério objectivo a atingir durante o treino. Por outro lado, a experimentação parece ser a forma de obter os melhores parâmetros, embora as diferenças sejam muitas vezes de pormenor. A determinação dos parâmetros de aprendizagem é normalmente empírica, baseada na experiência do utilizador e métodos de tentativa e erro (Henriques e Bação, 2004).

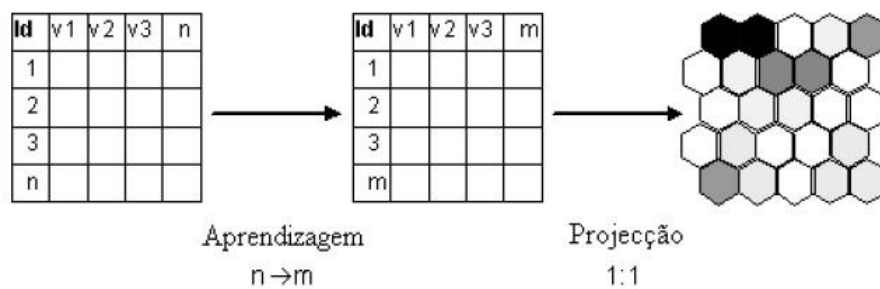


Figura 31 – Ilustração dos princípios do SOM (Henriques e Bação, 2004).

Os parâmetros fundamentais do *SOM* são a taxa de aprendizagem e a função de vizinhança. A taxa de aprendizagem (ta) define o grau de actualização dos pesos dos neurónios entre épocas (ciclos de apresentação de todos os indivíduos à rede neuronal), a qual assume um valor definido pelo utilizador no intervalo $[0,1]$ e que deve decrescer com o tempo (t) até um valor próximo de zero. A função de vizinhança $V(t)$ corresponde ao número de neurónios vizinhos a um neurónio cujos pesos são actualizados, sofrendo também actualização dos seus respectivos pesos. Esta função assume um valor definido no instante zero, $V(0)$, o qual deve ser apropriado ao objectivo de maximizar a qualidade da formação do mapa, e decresce também para zero ao longo do processo de aprendizagem.

5.5.3.2 O problema “*p-median*” e a construção de regiões

O problema da localização óptima de infra-estruturas baseia-se na procura da localização para um qualquer tipo de infra-estrutura, num espaço determinado, de tal forma que um ou vários critérios sejam optimizados. Neste contexto, o problema *p-median* é, geralmente, conhecido como o problema de localizar P infra-estruturas relativamente a um conjunto de clientes, de forma a minimizar a soma do caminho mais curto entre os “clientes” e as infra-estruturas. Uma tarefa potencialmente relacionada com o problema *p-median* tem a ver com a construção de m regiões a partir de n áreas, sendo que $n > m$, de acordo com um critério pré-determinado (Henriques e Bação, 2004).

O problema *p-median* é aqui encarado como ponto de partida para a resolução do problema de encontrar localizações para serviços de urgência, entendendo-se que se devem distribuir no território de forma equilibrada, no sentido de otimizar a quantidade de população residente coberta por cada um destes serviços.

O *p-median* consiste na localização de vários nós que minimizam a soma das distâncias entre cada nó e os seus “clientes”. Para transformar o *p-median* num problema de construção de regiões, basta considerar os clientes como representantes da população e os nós como representantes das regiões. A metodologia seguida neste estudo é muito similar à seguida por Henriques e Bação (2004), embora com algumas diferenças relacionadas com o presente caso de estudo. Assim sendo, apresenta-se aqui, para concluir esta secção, uma breve adaptação ao caso de estudo (desenvolvido em detalhe na secção seguinte).

O primeiro passo consiste em gerar pontos de forma aleatória para cada unidade espacial de população (neste estudo a secção estatística), baseados na respectiva população residente (no presente caso de estudo foi também ensaiada outra variável, referida na secção seguinte). A densidade dos pontos gerados é, assim, directamente proporcional ao valor associado a cada secção estatística; quanto maior este for, tanto maior será o número de pontos representativos. O conjunto de pontos gerados é utilizado para treinar um *SOM*, sendo que, no final do treino, cada neurónio do *SOM* representa um subconjunto dos pontos originais, na medida em que a cada um dos pontos é atribuído um neurónio vencedor. O posicionamento dos neurónios no final do treino é, assim, determinado pela quantidade de pontos que cada uma das diferentes secções estatísticas possui. Os neurónios são distribuídos de forma a minimizar a distância entre cada um deles e os pontos que representam. Por último, o passo seguinte, que consistiria na construção de regiões, não foi por enquanto concretizado, remetendo-se esta discussão para a última secção.

5.5.3.3 Abordagem metodológica

A área de estudo engloba todo a área do Distrito do Porto (figura 32). Esta escolha foi essencialmente ditada pela disponibilidade de dados, os quais são essencialmente de três tipos; por um lado, foram empregues registos da procura diária em serviços de urgência de hospitais e centros de saúde em Portugal Continental no ano de 2007, recolhidos pela DGS. Estes foram recolhidos e processados em ambiente *Microsoft Access 2003*, utilizando para o efeito a ligação ODBC estabelecida com a base de dados *Oracle* do SSESF.

Também a partir do SSESF, foram recolhidas as localizações pontuais dos serviços de urgência da área de estudo (suporte *shapefile*), as quais foram ainda

verificadas, e em alguns casos editadas, através da migração destas localizações para ambiente *Google Earth*. Foram ainda utilizados dados do INE, referentes à população residente por secção estatística (CENSOS 2001).

Quanto à base cartográfica deste estudo, consiste nos limites das secções estatísticas contidas na área escolhida, referidos à Base Geográfica de Referenciação Espacial (BGRI) do INE, a qual, devido à necessidade de interacção com a aplicação *Google Earth*, foi convertida do seu sistema nativo para o sistema de coordenadas *WGS84 (World Geodetic System 1984)*, utilizado consistentemente ao longo deste estudo. A maior parte das tarefas de preparação e análise de dados geográficos foram realizadas em ambiente *ESRI ArcGIS (ArcView) 9.2* e *ArcView 3.3*, com excepção da análise de localizações com recurso a redes neurais, efectuada através da aplicação *soMGis*, desenvolvida pelo Eng. Roberto Henriques e gentilmente cedida pelo mesmo, permitindo posteriormente exportar os resultados para suporte *shapefile*.

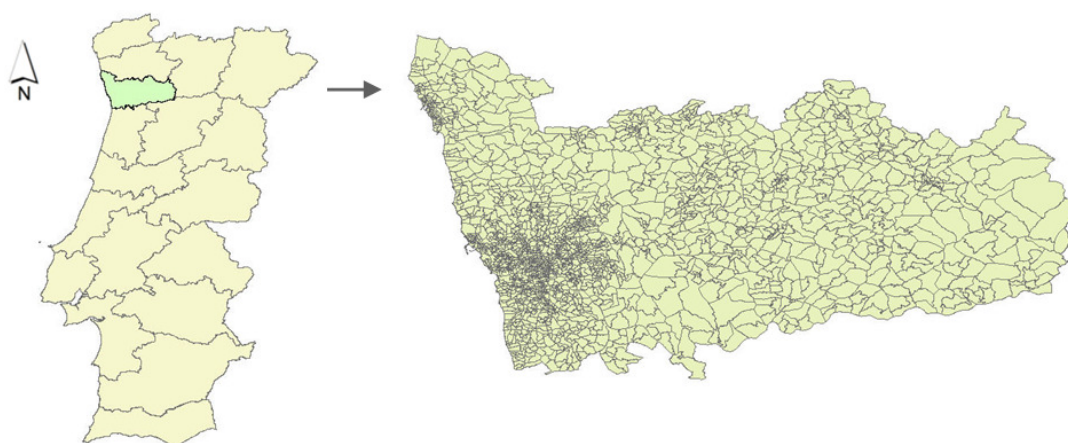


Figura 32 – Área de estudo.

Alguns pressupostos de base nortearam o desenrolar deste estudo. Definiu-se o ano de 2007 com referência, sendo o ano completo mais recente, para o qual estão disponíveis dados de procura nas urgências. Por outro lado, interessava escolher uma área de estudo na qual ocorresse o mínimo de alterações no decorrer do ano em estudo, no sentido de evitar a alteração da coerência das condições de estudo, o que levou à escolha do Distrito do Porto. No caso dos serviços de urgência, entende-se que o espaço é contínuo, uma vez que, ao contrário da procura de consultas de especialidade, os serviços de urgência de uma determinada área são

facilmente acessíveis a pessoas residentes em outras áreas do país; a limitação territorial existente imposta neste estudo teve também a ver com o tempo necessário para a recolha, preparação e análise dos dados, bem como para a verificação de resultados.

Foram seguidos dois rumos distintos para a resolução do problema *p-median* no presente caso de estudo.

A primeira linha de análise consistiu na utilização dos totais da procura de serviços de urgência como variável independente para a proposta de realocização. Estes valores de procura englobam, nesta fase de evolução do estudo, todas as idas a serviços de urgência de hospitais e centros de saúde, e por todos os motivos (e.g. acidente, doença, etc.), uma vez que o critério adoptado consiste em considerar que cada serviço deve servir o máximo possível de população residente, independentemente de outras considerações que serão deixadas para fases posteriores de aperfeiçoamento do modelo. Todavia, uma vez que os valores de procura disponíveis se encontram associados às unidades de prestação de serviços de saúde e não (pelo menos directamente) a unidades territoriais, foi necessário encontrar um critério para distribuir estes valores pelas secções estatísticas. Para tal, recorreu-se a áreas de influência delineadas pelo Ministério da Saúde, as quais indicam, para cada unidade de prestação de serviços de saúde, as respectivas freguesias de influência. Assim, associou-se aos limites de freguesia o código identificador de cada hospital ou centro de saúde da área de estudo com serviço de urgência a funcionar, em cuja área de influência a freguesia se insere. Calculou-se o total de população residente por área de influência, através da sobreposição dos centros das secções estatísticas com as áreas de influência assim obtidas, e também a proporção de população residente por secção estatística face ao total previamente encontrado por área de influência. Por último, através da multiplicação do total de procura por serviço de urgência por esta proporção, repartiu-se o valor de procura por todas as subsecções estatísticas de cada área de influência. Os valores assim obtidos serviram de base para a geração aleatória de pontos por secção estatística apresentados à rede neuronal.

A segunda linha de análise, muito mais simples, consistiu na geração dos pontos aleatórios apenas com base nos valores de população residente por secção estatística.

O passo seguinte consistiu na utilização do *soMGis* para encontrar novas localizações para os serviços de urgência em cada um dos cenários delineados. O

processo iniciou-se, em ambas as situações, pelo carregamento na aplicação de um tema vectorial (suporte *shapefile*) contendo os limites das secções estatísticas com um valor associado que corresponde à variável da qual se pretende fazer depender a reorganização dos nós (menu Temas → Adicionar). Foram de seguida gerados pontos aleatórios dentro da área de cada secção estatística, cuja quantidade é proporcional ao valor da variável, associado a cada uma das secções (menu SOM → Kohonen SOM). O soMGis permite escolher um limiar máximo de pontos a gerar antes de se dar início ao processo com o botão *Populate*, sendo que em ambos os casos, e após várias tentativas, foi escolhido um limiar de 50 pontos. A figura 33 apresenta um aspecto do processo após a geração dos pontos aleatórios. O botão “*Run SOM*” inicia o “painel de controlo” do SOM propriamente dito, no qual se introduzem os respectivos parâmetros.

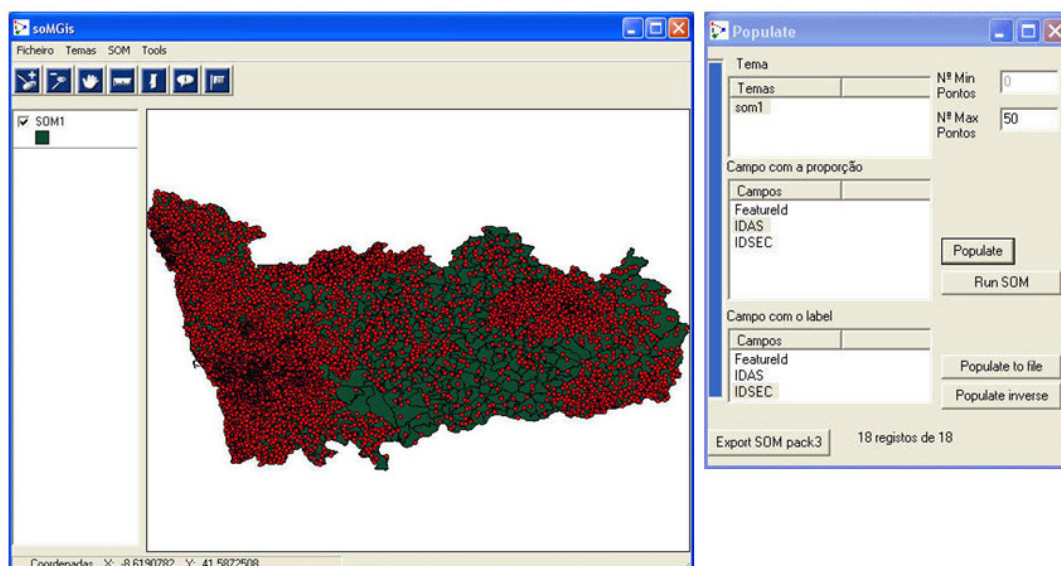


Figura 33 – Opções iniciais da aplicação.

A figura 34 ilustra o resultado final de uma das experiências, já com os neurónios gerados, bem como a parametrização utilizada. Após várias tentativas, a análise foi realizada, para ambas as linhas de análise, começando com uma taxa de aprendizagem de 0,6 (Carrão *et al*, 2006, referem que este parâmetro deve ser regulado entre 0,1 e 0,7) e uma matriz unidimensional com 43 neurónios (os quais correspondem ao número de serviços de urgência em funcionamento na área de estudo durante o ano de 2007), a qual é apontada por Henriques e Bação (2004), como uma melhor opção na tentativa de abordagem do problema p-median em

questão do que a utilização de duas dimensões. Quanto ao valor de vizinhança, foi regulado em 15 neurónios, o que representa aproximadamente metade dos desejados. Foram efectuadas duas experiências para cada um dos cenários em análise, mantendo os parâmetros acima referidos, mas usando dois valores distintos de “*momentum*” (0,9999 / 0,99999).

No sentido de tornar possível a comparação entre a situação actual e aquelas delineadas pelo SOM, foram gerados *polígonos de thiessen* que permitiram dividir a área de estudo com base nas localizações actuais (figura 35) e também naquelas encontradas pelo SOM (figura 36), repartindo-a geometricamente entre localizações.

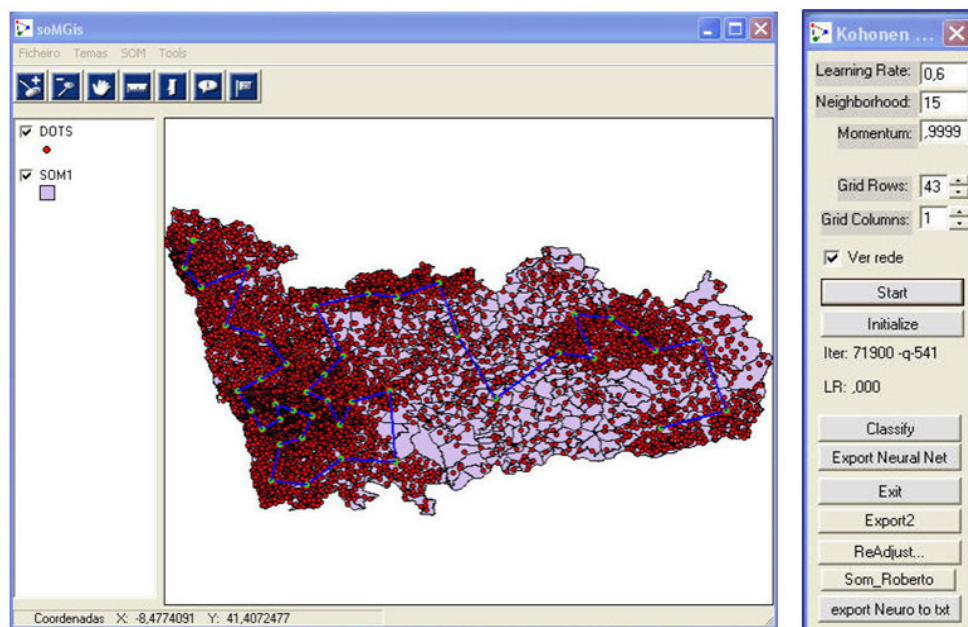


Figura 34 – Final do treino da rede neuronal e parametrização.

Estes polígonos foram então sobrepostos a pontos representativos da localização das secções estatísticas, previamente extraídos incluindo os respectivos valores de população residente, e tendo o cuidado de não utilizar centróides (por representarem centros geométricos, pode acontecer, dependente do grau de excentricidade do polígono, ficarem numa posição fora deste) mas sim *label points* (que se posicionam garantidamente dentro da área de cada secção).

Uma vez associada a cada secção estatística o polígono de thiessen a que pertence, procedeu-se à agregação (*summarize*) dos valores de população das

secções estatísticas para cada um destes polígonos, realizada em ambiente ESRI *ArcGIS*, na tabela resultante da operação de sobreposição.

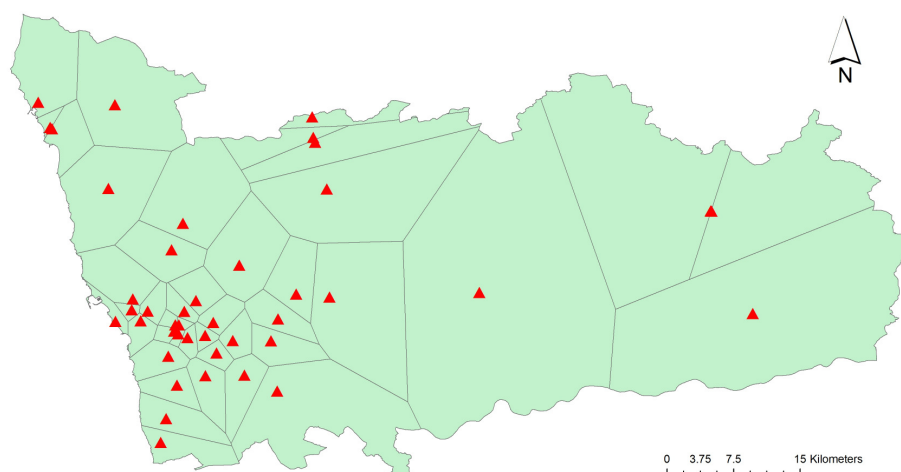


Figura 35 – Localizações actuais de serviços de urgência e polígonos de *thiessen*.

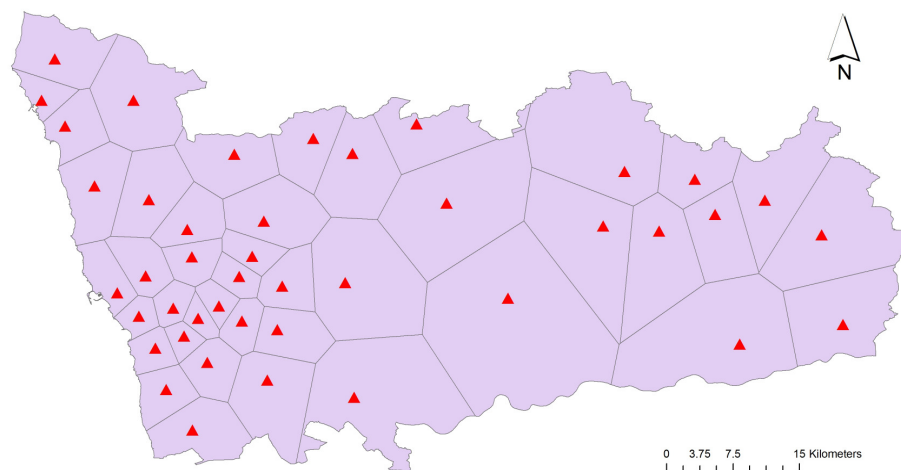


Figura 36 – Localizações alternativas para serviços de urgência e polígonos de *thiessen*.

5.5.3.4 Discussão de resultados

Os resultados obtidos apontam para uma melhoria da situação em qualquer um dos cenários gerados com recurso ao *SOM*, face à situação real em 2007 no Distrito do Porto, a qual apresenta um valor médio de 761 habitantes servidos por secção estatística. A tabela 3 contém os valores finais obtidos nas quatro experiências realizadas. O aumento de casas decimais no parâmetro “*momentum*” parece ter piorado um pouco os resultados em ambas as situações.

O melhor valor (837 habitantes servidos por secção estatística) foi alcançado com um parâmetro de “*momentum*” de 0,9999 e no cenário mais simples, no qual se utilizou directamente o valor da população residente por secção estatística.

Por outro lado, o recurso à variável composta referente à procura em serviços de urgência apresentou piores resultados do que a utilização isolada da população residente. Esta situação pode ter a ver com a configuração das áreas de influência dos serviços de saúde em estudo, as quais obedecem aos limites administrativos, o que vem possivelmente, ao invés do pretendido, piorar a variável composta produzida.

Variável utilizada	<i>Momentum</i>	
	0,9999	0,99999
População residente	837	833
Procura * proporção da População Residente	809	800

Tabela 3 – Resultados dos cenários de realocização ensaiados (habitantes servidos por secção estatística).

Também uma parte da procura destes serviços tem ainda ver com pessoas não residentes na área de estudo, o que diminui igualmente o poder desta variável, uma vez que a análise está, para já, confinada ao Distrito do Porto.

5.5.3.5 Notas finais e recomendações

Este estudo constitui um primeiro ensaio da utilização da metodologia *SOM* no presente contexto. Todavia, apesar dos resultados obtidos serem, para já, embrionários, o facto de aparentarem melhorar a configuração territorial da rede de urgências encoraja ao aprofundamento da metodologia. Desta forma, prevêem-se várias linhas futuras de experimentação.

Em primeiro lugar, impõe-se mitigar algumas fontes potenciais de enviesamento, sendo importante garantir a recepção de dados de todas as instituições prestadoras de serviços de saúde com serviços de urgência em funcionamento, o que ainda não é uma realidade para todo o país, por vários motivos. Resolvidos estes problemas, o estudo pode ser estendido à totalidade do território nacional, o que evitará efeitos de fronteira num espaço que, como já referido, é contínuo.

Delinear, após esta primeira experiência, modelos separados para hospitais e centros de saúde poderá trazer vantagens, uma vez que ambos possuem diferentes capacidades para receber pacientes, especialidades e impacto no território

Outra possibilidade, assumindo a continuação da experimentação com recurso ao SOM, reside em encontrar uma variável composta que represente eficazmente cada subsecção, conjugando num só vários valores relevantes, para gerar os pontos aleatórios. Por outro lado, o recurso aos valores de procura por serviço não pode ser considerado ainda encerrado, na medida em que carece, ao invés, de um critério melhor do que o empregue, que permita distribuir estes valores pelas subsecções de forma a evitar enviesamentos.

O esforço de melhoramento dos parâmetros do SOM é uma questão que merece ainda a reserva de algum tempo na agenda.

O cálculo de valores de comparação entre configurações territoriais com recurso a *polígonos de thiessen* poderá ainda ser substituído pelo cálculo de áreas de serviço mutuamente exclusivas para cada nova configuração, baseado em análise de redes efectuada tendo como base a rede rodoviária portuguesa, o que trará, em princípio, uma comparação mais realista.

Por último, o passo seguinte, que consistiria na construção de regiões, não foi por enquanto concretizado. A construção das regiões pode ser feita pelo *soMGis*, associando a cada neurónio as secções estatísticas que se encontram mais próximo desse neurónio do que de qualquer outro. A associação é feita com base nas distâncias medidas entre os neurónios e as fronteiras das diferentes secções. Ou seja, para cada secção é determinado o neurónio que se encontra mais próximo, e a aglomeração de secções que partilham o mesmo neurónio forma as várias regiões. Esta metodologia poderia vir a servir de base para a geração de novas áreas de influência para serviços de urgência, após a delineação de uma configuração territorial alternativa que possua reconhecidamente um grau de eficácia superior ao da actual.

5.6 Conclusões e desenvolvimentos futuros do SESP

Para além da estrutura do sistema e das funcionalidades que já se encontram operacionais, estão previstos ainda desenvolvimentos para o SESP, cuja implementação será efectuada gradualmente.

Uma das funcionalidades previstas consiste em disponibilizar ao utilizador a possibilidade de recorrer ao sistema para introduzir dados epidemiológicos remotamente, tais como a localização pontual do foco de um surto ou epidemia, uma ocorrência de carácter ambiental (e.g. um poço contaminado) ou a localização de uma nova infra-estrutura de prestação de cuidados de saúde. A figura 37 apresenta uma simulação desta funcionalidade.

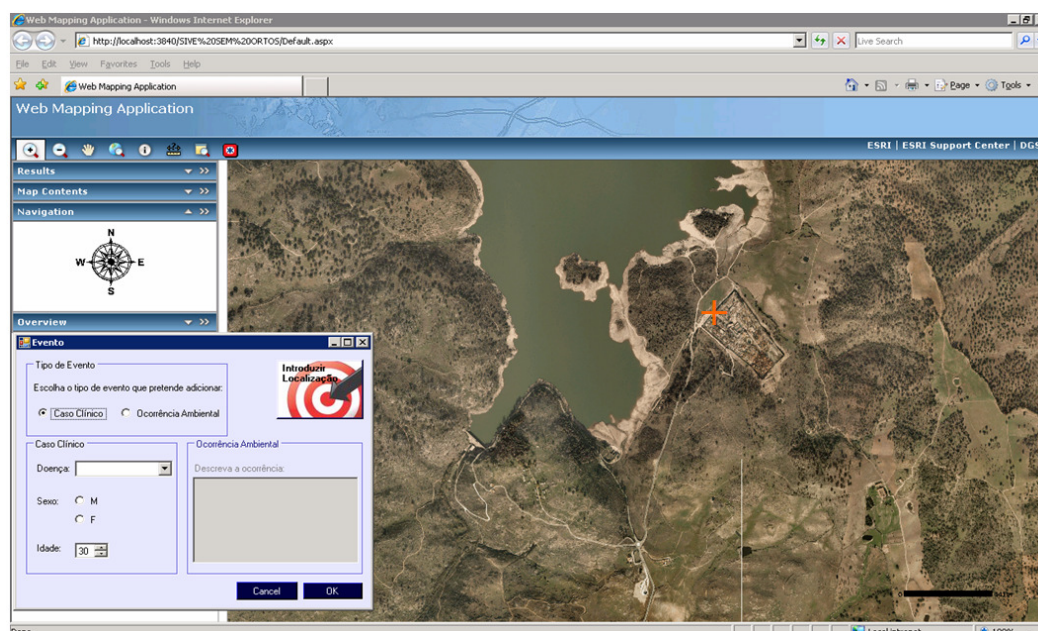


Figura 37 – Georrefenciação de um foco epidémico detectado num parque de campismo.

Assim, a total cobertura do território continental com ortofoto-imagens digitais, associada a outros tipos de informação geográfica (e.g. eixos de via e respectiva designação, limites administrativos, etc.), auxilia o utilizador na localização do evento que deseja georreferenciar; de seguida, pressionando com o rato num simples botão, surge uma caixa de diálogo que permite introduzir a localização e associar-lhe os atributos necessários. No final, as coordenadas e os atributos da localização inserida são automaticamente armazenados numa tabela da base de dados *Oracle*, para utilização futura.

Este tipo de recolha de dados pode ser muito útil para certos tipos de análise espacial que, pela sua complexidade técnica ou carácter reservado, apenas são levadas a cabo na DGS em situações específicas. Pode ainda ser utilizado para fazer o levantamento da localização de infra-estruturas de prestação de serviços de saúde existentes no território, permitindo manter actualizados os respectivos temas.

Pretende-se também implementar a emissão de alertas e a produção de corredores epidémicos semanais em escalas de análise maiores do que a nacional, a qual se pode revelar, por vezes, insuficiente. Prevê-se ainda que a emissão de alertas seja efectuada directamente para telefones móveis, em texto do serviço de mensagens curtas (*short message service* - *SMS*), podendo esta funcionalidade ser programada directamente na base de dados *Oracle*.

Por questões relacionadas com a análise estatística envolvida na produção dos corredores epidémicos, a opção seguinte será pelo nível distrital, prevendo-se ainda a produção de mapas temáticos, automaticamente actualizados, ilustrando o nível de alerta diário. A figura 38 apresenta uma simulação desta funcionalidade, com base na procura de urgências no pico da epidemia sazonal de gripe de 2007, atingido a 5 de Fevereiro.

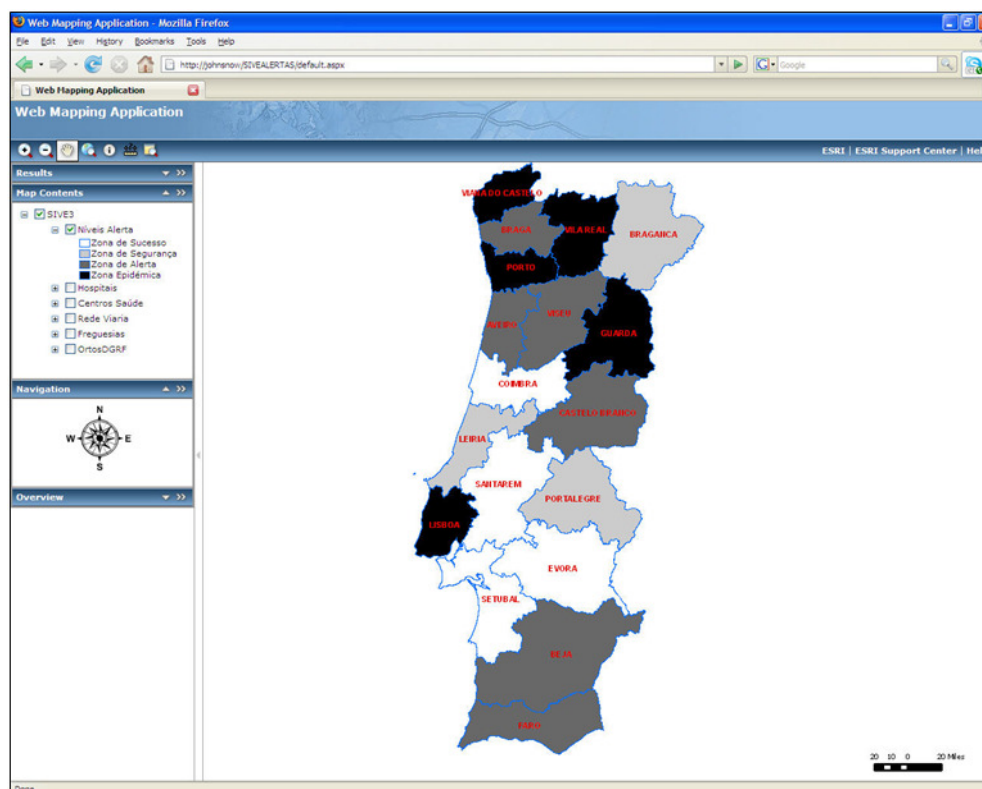


Figura 38 – Níveis de alerta distritais, 05-02-2007.

6 CONCLUSÕES

Impõem-se aqui alguns comentários de encerramento, relacionados com a concepção, implementação e perspectivas para o sistema aqui em foco, mas também com o contexto em que as tarefas se desenrolaram.

6.1 Contexto de desenvolvimento da aplicação

Tendo em conta que o processo aqui descrito se iniciou em princípios do ano de 2006, parece ter-se verificado um crescente interesse e aposta nas aplicações de informação geográfica em Saúde Pública, e uma progressão dos sistemas de informação para a saúde em geral, desde então. Tal como referido no segundo capítulo, a realização de um simpósio subordinado a este tema em Portugal é, sem dúvida, um factor muito positivo. Esta tendência parece verificar-se de igual modo ao nível internacional, bastando consultar o número e diversidade de temas das publicações surgidas no *International Journal of Health Geographics* desde o ano de 2006 até à data de publicação desta dissertação.

6.2 Desenvolvimento e utilidade do SESP

Acerca do SESP, este é um sistema em contínuo desenvolvimento, tendo já trazido benefícios no apoio à decisão em Saúde Pública em várias situações. À data da escrita deste texto, boa parte da arquitectura inicialmente prevista já se encontrava completa e operacional, subsistindo todavia algumas ferramentas por implementar, já referidas no capítulo anterior, algumas questões por resolver, discutidas mais à frente, e perspectivas para possíveis desenvolvimentos futuros.

No que respeita aos objectivos de vigilância epidemiológica do desenvolvimento proposto, é possível afirmar que a evolução foi positiva, na medida em que os pontos essenciais foram cumpridos:

- O protocolo de tarefas manuais anteriormente efectuadas numa base diária foi totalmente substituído por mecanismos de execução automática sem intervenção humana;

- O sistema actualmente em vigor constitui já uma plataforma de integração, agregando tecnologias muito diversas a funcionar em conjunto, e possuindo armazenamento e processamento simultâneo de dados alfanuméricos e geográficos, ao invés do que acontecia antes;
- Todas as tarefas diárias de armazenamento e processamento de dados alfanuméricos relativos à procura de serviços de urgência se encontram presentemente centradas numa aplicação que possui as capacidades e arquitectura de um verdadeiro SGBD relacional (*Oracle 10g*), com vantagens ao nível da velocidade de processamento, capacidade de agendamento de tarefas, robustez, segurança e potencial de escalabilidade futura do sistema;
- O sítio na Internet anteriormente existente foi substituído por um acesso que apresenta todos os resultados do sistema (alfanuméricos, gráficos e geográficos), para além de se apresentar muito mais amigável, intuitivo e interactivo com o utilizador.

São várias as vantagens desta infra-estrutura. Possibilita a partilha de conhecimento e informação aos utilizadores finais e vice-versa, no sentido de assegurar uma apreciação mais frequente e abrangente de situações anómalas em Saúde Pública, graças à disponibilização dos seus resultados na Internet. É utilizada tanto na monitorização de rotina em Saúde Pública, servindo como plataforma de armazenamento de dados (*Data Warehouse*) que são preciosos para a realização de estudos epidemiológicos de fundo, como na vigilância de síndromes, permitindo detectar surtos e emitir alertas. É de salientar o seu baixo custo para a maior parte dos utilizadores em termos de *hardware*, *software* e formação, pois apenas necessitam de possuir um computador e uma ligação à Internet, não obrigando a qualquer tipo de conhecimento de ferramentas SIG ou de consultas especializadas à bases de dados.

6.3 Limitações e melhoramentos para o SESP

Subsistem ainda algumas limitações na arquitectura dos sistemas SINUS e SONHO, nas comunicações com estes e ao nível da infra-estrutura montada

localmente na DGS, o que leva a que o sistema e os seus resultados tenham de ser monitorizados permanentemente.

Assim, parte das bases de dados locais das instituições de saúde em vigilância foram implementadas há já bastante tempo, e algumas correm ainda versões de *Oracle* bastante desactualizadas, caso da versão 6, para as quais não é possível estabelecer comunicações com versões superiores a *Oracle 9i*, razão pela qual esta versão pode ser presentemente encontrada no servidor de recolha. Esta máquina constitui, por si só, um problema e um estrangulamento do sistema, por ser antiga, lenta e dada a falhas físicas, estando já planeada a sua substituição.

Outra situação problemática reside nas modificações por vezes introduzidas por parte dos administradores das bases de dados locais das unidades hospitalares, que levam à interrupção temporária da transferência de dados a partir da respectiva unidade de saúde, sendo necessário um maior esforço de coordenação para evitar estas situações.

O facto dos sistemas SINUS e SONHO funcionarem em separado leva a existam diferenças nos dados que chegam de cada um deles. Um exemplo pode ser observado ao nível do motivo da ida à urgência em hospitais e centros de saúde, pois os respectivos registos apresentam um código numérico diferente para o mesmo motivo (e.g. doença).

Uma limitação realmente premente tem a ver com a ausência de registo preciso do diagnóstico associado a cada episódio de procura de serviços de urgência, sendo possível saber apenas o *motivo* em geral, tal como doença, acidente, entre outros. Existem já sistemas que registam também o diagnóstico, tal como a aplicação *GeoMedStat* (Li *et al*, 2006), que codifica cada episódio de procura com recurso à Classificação Internacional de Doenças, nona revisão (ICD-9), com resultados muito interessantes na vigilância e detecção de surtos.

Outra questão prende-se com o facto do SESP não receber ainda dados de todas as unidades de saúde com serviços de urgência do país, sendo desejável que estas redes se encontrassem totalmente cobertas, proporcionando uma percepção mais abrangente e real da evolução diária da situação de Saúde Pública em Portugal.

Estas situações chamam a atenção para a necessidade de uma revisão destes sistemas, que implicaria não apenas questões estruturais (ao nível do aparato de *hardware* e *software*), mas também questões organizacionais, tais como a adopção de um modelo conceptual (ou *ontologia*), transversal a todo o desenvolvimento, no

sentido de facilitar a integração de dados e levar a que todos os participantes envolvidos partilhem um mesmo diálogo e vocabulário no decurso das tarefas de implementação e de funcionamento corrente. A este nível, Boulos (2004), refere um conjunto de directivas para o desenvolvimento de sistemas de vigilância epidemiológica que poderão ser úteis como ponto de partida.

6.4 Próximos horizontes

Ultrapassados estes problemas (pelo menos os mais importantes, como a ausência de registo do diagnóstico preciso), a meta seguinte consistirá em evoluir para formas de análise mais complexas (e completas) do que as actualmente apresentadas.

Assim, na senda de tornar mais completa a actividade do sistema, é fundamental estender as tarefas de vigilância epidemiológica do SSES, podendo este passar a monitorizar muito mais do que a procura de serviços de urgência. De facto, o sistema possui um potencial de escalabilidade que possibilita efectuar a vigilância dedicada de várias doenças ao mesmo tempo, através da mesma *interface*, pelo que passariam a existir, no fundo, vários subsistemas dentro do SSES (e.g. vigilância da tuberculose, doenças de declaração obrigatória, e outras cuja vigilância se possa considerar premente). Esta perspectiva encontra-se presentemente limitada pelos dados recebidos, existindo já capacidade tecnológica e experiência para apostar na expansão do sistema.

Para além da integração de dados que indiciem directamente as ocorrências, outra perspectiva reside na inclusão de dados que o fazem indirectamente. Assim, no sentido de reforçar os actuais mecanismos de apoio à decisão e de vigilância de síndromes, a DGS considera ainda a inclusão de outros dados alfanuméricos, provenientes de *fontes-sentinelas*: inquéritos telefónicos realizados a intervalos regulares sobre um conjunto de agregados familiares escolhidos aleatoriamente com o objectivo de detectar sinais e sintomas que indiciem uma determinada doença (e.g. gripe) e formulários de monitorização preenchidos por empresas e/ou escolas para detectar flutuações anómalas no absentismo que possam indiciar o início de um surto. Estes formulários constituem, só por si, um exemplo interessante de aplicação de um sistema de informação para a saúde, uma vez que são disponibilizados e preenchidos através de um sítio na Internet próprio numa área de acesso reservado do sítio na Internet da DGS (tal como o SSES), sendo os seus

dados também armazenados para dar apoio a vários estudos de fundo. Toda a gestão e preparação destes formulários são também efectuadas, através de um sítio na Internet dedicado, que constitui uma *interface* de administração, conferindo a esta ferramenta uma grande flexibilidade.

Finalmente, existem já na literatura consultada para esta dissertação, um conjunto de técnicas e metodologias inovadoras muito apetecíveis, que podem ser integradas e operacionalizadas no âmbito do SESP, podendo vir a constituir linhas de desenvolvimento futuro para o sistema.

Como ilustrado no capítulo quinto, o sistema recorre presentemente a *mapas coropletos* para ilustrar a procura distrital de serviços de urgência. Assumindo que passaria a dispor de registos dos diagnósticos, seria possível cartografar directamente a incidência ou prevalência de diferentes doenças também através de mapas coropletos. Mas estes mapas, não obstante a sua utilidade, possuem várias limitações, tais como a excessiva simplificação visual (os valores aparecem uniformemente distribuídos dentro de cada polígono), e a variação abrupta dos valores nas fronteiras dos polígonos enquanto na realidade a incidência das doenças apresenta uma distribuição contínua no espaço, entre outras (Boulos, 2004). Para ultrapassar estas limitações, o autor refere que a análise dos dados de saúde se revela mais eficaz quando são utilizados métodos que assumem que a incidência das doenças é contínua no espaço, indicando várias técnicas para o efeito. Na eventualidade destes métodos virem a ser implementados ao nível do SESP, constituiriam um reforço do sistema, coexistindo com as actuais escalas e formas de análise.

Um ponto de partida possível para todos estes métodos, seria agregar os dados ao nível da mínima desagregação territorial possível no sistema, recorrendo à *residência* das pessoas atendidas, que é recolhida ao nível da Freguesia. A partir daqui, tornar-se-ia possível a aplicação de vários métodos, numa base diária, ou mesmo em *tempo real*, caso a infra-estrutura do SESP viesse a suportar este tipo de transferência e actualização de dados, existente já em alguns sistemas, notavelmente no sistema *RODS* (Boulos, 2004).

Uma alternativa aos mapas coropletos seria, então, a apresentação dos resultados agregados em áreas geográficas menores, através da preparação de *mapas suavizados* (*smoothed maps*), que apresentam os resultados como superfícies de variação suave.

Existem várias maneiras de produzir estes mapas, sendo que uma das mais intuitivas consiste na utilização de um *filtro espacial*, que pode ser aplicado a dados pontuais ou areais (neste caso seriam os polígonos representativos dos limites das freguesias).

Uma técnica possível consiste em considerar que a incidência verificada numa determinada localização corresponde à incidência observada até uma distância fixa dessa mesma localização. Os círculos associados aos pontos correspondentes às diversas localizações sobrepõem-se, sendo recalculada a incidência em cada localização, seguida do cálculo de isolinhas baseadas nos pontos representativos das localizações, sendo assim obtido um mapa suavizado. Um exemplo deste tipo de cartografia foi desenvolvido e disponibilizado num sítio na Internet por Beyer (2006), ilustrando a variação espacial da incidência de cancro colorectal no estado norte-americano do Iowa, entre 1998 e 2003.

Boulos (2004) refere ainda outras técnicas para obtenção de cartografias representativas da variação da incidência que assumem que esta é contínua no espaço, tais como o recurso a técnicas de interpolação geostatística como o *kriging*, a produção de mapas de superfícies de tendência (*trend surface maps*), ou o recurso a métodos de inferência *bayesiana*. Estas representações podem ser muito úteis em vigilância epidemiológica, pois permitem avaliar o estado da arte em termos de incidência das doenças vigiadas de uma forma mais próxima da realidade, facilitando também a detecção de surtos.

Todavia, estes métodos de cartografar ocorrências, por mais próximos da realidade que se encontrem, permitem apenas uma avaliação visual do padrão de distribuição espacial da incidência, sendo a detecção de surtos efectuada também visualmente, processo que se revela, em última análise, propenso ao erro humano. Uma forma de ultrapassar esta limitação seria dotar o SESP da capacidade de levar a cabo análise de *clusters*, efectuada automaticamente numa base diária ou em tempo real. Apesar deste tipo de análise ser essencialmente exploratório, a sua aplicação pode ser útil como forma de detecção precoce de surtos, mesmo que tenham de ser confirmados (e.g. através da validação por peritos). Existem vários métodos disponíveis, como já referido no segundo capítulo, mas a escolha do mais apropriado pode revelar-se uma questão delicada. A título de exemplo, Rogerson (1997) discute uma metodologia possível para usos de carácter proactivo, tal como em tarefas de vigilância epidemiológica. Estes métodos possuem vantagens, por serem explicitamente dedicados à detecção destes fenómenos, e por avaliarem

estatisticamente a validade dos *clusters* detectados, o que vem conferir uma base científica mais sólida às análises efectuadas, além de facilitar a sua possível implementação no SESP como mecanismo automático. Elliot e Wartenberg (2004) salientam a utilidade destes métodos nos sistemas de vigilância de carácter proactivo, mais do que na realização de estudos *ad-hoc* numa base local para identificar e priorizar casos em excessos para posterior investigação. Rogerson (1997) refere que este tipo de utilização é relativamente rara, mas existem já exemplos de sistemas de vigilância que os incorporam. Nesta perspectiva, um caso de estudo que poderia constituir inspiração para a implementação da análise de *clusters* no SESP é o sistema *EpiScanGIS*, já apresentado no terceiro capítulo.

Para além dos possíveis melhoramentos e rumos futuros relacionados com a vigilância epidemiológica diária ou semanal, restam as aplicações que tiram partido da infra-estrutura e funcionalidades existentes, não sendo, todavia, efectuadas numa base regular. A este nível, as expectativas apontam para o prosseguimento das actividades e estudos referidos no capítulo cinco, na óptica da continuidade, mas também do aperfeiçoamento contínuo das metodologias empregues.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERT, D., GESSLER, W. e LEVERGOOD, B. (Ed.s), 2000, *Spatial analysis, GIS and Remote Sensing applications in the Health Sciences* (Chelsea: Ann Arbor Press).
- ALEA, 1999, *Acção Local Estatística Aplicada* (URL: <http://www.alea.pt/html/actual/html/act06.html>, consulta em 5-4-2008).
- ALVES, J., LOPES, P., e LOURENÇO, P., 2001, Utilização de Detecção Remota para a Determinação de Zonas de Risco de Contaminação por Malária e a Influência das Alterações Climáticas Globais na sua Expansão – Região do Algarve como Caso de Estudo. Apresentado no *Seminário de Verão do Grupo de Análise de Sistemas Ambientais (GASA)*. (URL: http://air.dcea.fct.unl.pt/pedro_garrett/papers/RSMalaria.pdf, consulta em 10-4-2008).
- BALKÁNYI, L., 2007, Terminology services - an example of knowledge management in public health. *Euro Surveillance*, **12**(22), [Article ID: 3211] (URL: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=3211>, consulta em 11-6-2008).
- BEAGLEHOLE, R., BONITA, R., e KJELLSTRÖM, T., 1993, *Basic Epidemiology* (Geneva: World Health Organization).
- BEYER, K., 2003, *Colorectal Cancer Incidence Rate, 1998-2003* (URL: http://www.uiowa.edu/~qishlth/ICCCCMaps/index_Page446.htm, consulta em 12-6-2008).
- BIOMED CENTRAL, 1999, *Biomed Central* (URL: <http://www.biomedcentral.com/>, consulta em 7-7-2008).
- BLANTON, J. D., MANANGAN, A., MANANGAN, J., HANLON, C. A., SLATE, D e RUPPRECHT, C. E., 2006, Development of a GIS-based, real-time Internet mapping tool for rabies surveillance. *International Journal of Health*

- Geographics*, 5:47 (URL: <http://www.ij-healthgeographics.com/content/5/1/47>, consulta em 13-7-2008).
- BORTMAN, M., 1999, Elaboracion de corredores o canales endémicos mediante planillas de cálculo. *Revista Panamericana de Salud Pública*, 5(1), 8 pp.s (URL: <http://www.scielosp.org/pdf/rpsp/v5n1/5n1a1.pdf>, consulta em 5-5-2006).
- BOSCOE, F., WARD, M., e REYNOLDS, P., 2004, Current practices in spatial analysis of cancer data: data characteristics and data sources for geographic studies of cancer. *International Journal of Health Geographics*, 3:28 (URL: <http://www.ijhealthgeographics.com/content/3/1/28>, consulta em 10-10-2006).
- BOULOS, M., ROUDSARI, A., e CARSON E., 2001, Health Geomatics: an enabling suite of technologies in health and healthcare. *Journal of Biomedical Informatics*, 34, pp. 195–219 (URL: <http://www.idealibrary.com>, consulta em 10-10-2006).
- BOULOS, M., 2004, Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom. *International Journal of Health Geographics*, 3:1 (URL: <http://www.ijhealthgeographics.com/content/3/1/1>, consulta em 10-10-2006).
- BOULOS, M., 2004, Descriptive review of geographic mapping of severe acute respiratory syndrome (SARS) on the Internet. *International Journal of Health Geographics*, 3:2 (URL: <http://www.ijhealthgeographics.com/content/3/1/2>, consulta em 29-7-2008).
- BOULOS, M., 2005, On geography and medical journalology: a study of the geographical distribution of articles published in a leading medical informatics journal between 1999 and 2004. *International Journal of Health Geographics*, 4:7 (URL: <http://www.ij-healthgeographics.com/content/4/1/7>, consulta em 29-7-2008).

- CARRÃO, H., CAPÃO, L., BAÇÃO, F., e CAETANO, M., 2006, Meris based land cover classification with self-organizing maps: preliminary results. In *Proceedings of the 2nd EARSeL SIG Workshop on Land Use & Land Cover*, 28 – 30 September 2006 (Bonn, Germany), 8 pp.s (URL: <http://www.igeo.pt/gdr/pdf/Carrao2006c.pdf>, consulta em 3-4-2008).
- CARREIRA, M., OLIVEIRA, A., CARDOSO, C. e PLANTIER, T., 2007
Desenvolvimento do Sistema Integrado de Vigilância Epidemiológica da Direcção-Geral da Saúde [CD-ROM]. In *Actas do 1º Congresso de Saúde e Tecnologias de Informação Geográfica*, 31 de Maio - 1 de Junho, 2007, 13 pp.s.
- CARREIRA, M., ALARCÃO, V., OLIVEIRA, A., CARDOSO, C. e PLANTIER, T., 2007, Amostragem espacial na avaliação do nível de saúde e acessibilidade aos cuidados de saúde nas comunidades imigrantes [CD-ROM]. In *Actas do 1º Congresso de Saúde e Tecnologias de Informação Geográfica*, 31 de Maio - 1 de Junho, 2007, 9 pp.s.
- CDC, 2001, *Updated Guidelines for Evaluating Public Health Surveillance Systems* (URL: <http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm>, consulta em 7-7-2008).
- CDC, 2008, *What Is Epi Info* (URL: <http://www.cdc.gov/epiinfo/>, consulta em 1-6-2007).
- COLIN, P., 2005, *Dictionary of Medical Terms* (4ª ed.) (London: A & C Black).
- CORVALLIS, 2000, *Introduction to the Global Positioning System for GIS and TRAVERSE* (URL: <http://www.cmtinc.com/gpsbook/index.htm>, consulta em 10-5-2008).
- CROMLEY, E., e MCLAFFERTY, S., 2002, *GIS and Public Health* (New York: The Guilford Press).

- DAWES, C., BRYLA, B., JONHSON, J. C., e WEISHAM, M., 2005, *OCA: Oracle 10g Administration I, Study Guide* (San Francisco: Sybex).
- DGS, 2008, Página da Direcção Geral da Saúde, (URL: <http://www.dgs.pt/>, consulta em 19-7-2008).
- ELLIOT, P., WAKEFIELD, J. C., BEST, N. G., e BRIGGS, D. J. (Ed.s), 2000, *Spatial Epidemiology: methods and applications* (Oxford: Oxford University Press).
- ELLIOT, P., e WARTENBERG, D., 2004, Spatial Epidemiology: Current Approaches and Future Challenges. *Environmental Health Perspectives*, vol. 112, nº 9, pp. 998-1006 (URL: <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1247193&blobtype=pdf>, consulta em 10-7-2008).
- EPA, 2008, *Thesaurus of Terms Used in Microbial Risk Assessment* (URL: <http://www.epa.gov/waterscience/criteria/humanhealth/microbial/thesaurus/T57.html>, consulta em 22-7-2008).
- EPISCANGIS, 2005, *EpiScanGIS. Meningococcal Disease Surveillance in Germany* (URL: <http://episcangis.hygiene.uni-wuerzburg.de/Locale.do?language=en>, consulta em 8-7-2008).
- ESRI, 2006, *GIS Dictionary* (URL: <http://support.esri.com/index.cfm?fa=knowledgebase.gisDictionary.search&searchTerm=GIS>), consulta em 10-10-2006).
- FAO, 2008, *Emergency Prevention System (EMPRES) for Transboundary Animal and Plant Pests and Diseases* (URL: <http://www.fao.org/ag/againfo/programmes/en/empres/home.asp>, consulta em 22-5-2008).
- GATRELL, A., e LÖYTÖNEN, M. (Ed.s), 1998, *GIS and Health*, (London: Taylor & Francis).

- GILBERT, M., 2005, *Spatial epidemiology of highly pathogenic avian influenza (HPAI) in Thailand. Exploring the use of remote sensing for risk assessment.* (URL: http://www.fao.org/avianflu/conferences/rome_avian/documents/Spatial%20epidemiology%20of%20HPAI%20in%20Thailand.pdf, consulta em 5-3-2007).
- GOOVAERTS, P., GEBREAB, S., 2008, How does Poisson kriging compare to the popular BYM model for mapping disease risks? *International Journal of Health Geographics*, 7:6 (URL: <http://www.ijhealthgeographics.com/content/7/1/6>, consulta em 10-7-2008).
- GOSSELIN, P., LEBEL G., RIVEST S., e Douville-Fradet, M., 2005, The Integrated System for Public Health Monitoring of West Nile Virus (ISPHM-WNV): a real-time GIS for surveillance and decision-making. *International Journal of Health Geographics*, 4:21 (URL: <http://www.ij-healthgeographics.com/content/4/1/21>, consulta em 10-7-2008).
- HASSAN, A.N., KENAWY, M.A., KAMAL, H., ABDEL SATTAR, A. A., e SOWILEM, M. M., 2003, GIS-based prediction of malaria risk in Egypt. *La Revue de Santé de la Méditerranée orientale*, 9(4), pp. 548-558 (URL : <http://www.emro.who.int/publications/emhj/0904/PDF/07.pdf>, consulta em 10-10-2006).
- HEALTH CANADA, 2000, *Case definitions for diseases under national surveillance.* (URL: <http://www.phac-aspc.gc.ca/publicat/ccdr-rmtc/00vol26/26s3/index.html>, consulta em 10-7-2008).
- HENRIQUES, R., e BAÇÃO, F., 2004, soMGis: Uma Ferramenta para Construir Regiões. In *V Congresso da Geografia Portuguesa* (Campus de Azurém – Guimarães, 14-16 de Outubro), 20 pp.s (URL: http://www.isegi.unl.pt/ensino/docentes/fbacao/Henriques_VCongGeo2004.pdf, consulta em 3-4-2008).
- IJHG, 1999, *International Journal of Health Geographics* (URL: <http://www.ij-healthgeographics.com/home/>, consulta em 1-10-2006).

- JACQUEZ, G. M., e GREILING, D. A., 2003, Local clustering in breast, lung and colorectal cancer in Long Island, New York. *International Journal of Health Geographics* 2:3 ([URL:http://www.ijhealthgeographics.com/content/2/1/3](http://www.ijhealthgeographics.com/content/2/1/3), consulta em 10/10/2006).
- KHAN, O. (Ed.), 2003, *Geographic Information Systems and Health Applications* (Hershey: Idea Group Publishing).
- KOHONEN, T., 2001, *Self-Organizing Maps* (3ª ed.) (Springer).
- KULLDORFF, M., 1997, A Spatial Scan Statistic, *Communications in Statistics: Theory and Methods*, 26(6), pp. 1481-1496 (URL: <http://www.satscan.org/papers/k-cstm1997.pdf>, consulta em 10/7/2008).
- LAWSON, A., BIGGERI, A., BÖHNING, D., LESAFFRE, E., VIEL, J., BERTOLLINI, R. (Ed.s), 1999, *Disease Mapping and Risk Assessment for Public Health* (West Sussex: John Wiley & Sons).
- LAWSON, A. B., 2006, *Spatial methods in spatial epidemiology*, (2ª ed.) (West Sussex: John Wiley & Sons).
- LI, H., FARUQUE, F., WILLIAMS, W., e FINLEY, R., 2006, Real-Time Syndromic Surveillance. *ArcUser magazine*, vol. 9, nº1, pp. 17-19 (Redlands: EsriPress).
- LOPES, P., LOURENÇO, P., SOUSA, C., NOVO, T., RODRIGUES, J., ALMEIDA, A., e SEIXAS, J., 2005, Modelling Patterns of Mosquito Density Based on Remote Sensing Images. In *II International Conference and Exhibition on Geographic Information*, Estoril Congress Center, May 30 - June 2, 9 pp.s. (URL: http://air.dcea.fct.unl.pt/pedro_garrett/papers/paper177.pdf, consulta em 10-4-2008).
- MARTIN, V., Von DOBSCHUETZ S., LEMENACH, A., RASS, N., Schoustra W., e DeSimone, L., 2007, Early Warning, Database and Information Systems for Avian Influenza Surveillance. *Journal of Wildlife Diseases, Wildlife Disease Association* 43(3) Supplement 2007, pp. 71-76. (URL:

- http://www.jwildlifedis.org/cgi/reprint/43/3_Supplement/S71.pdf, consulta em 10-7-2008).
- NIAID, 2008, *Enteric Diseases* (URL: <http://www3.niaid.nih.gov/research/topics/enteric/>, consulta em 26-07-2008).
- NIELSEN, J., LORANGER, H., 2006, *Prioritizing Web Usability*. (Berkeley: New Riders).
- OMS, 2008, *The HealthMapper* (URL: http://www.who.int/health_mapping/tools/healthmapper/en/, consulta em 21-07-2007).
- ORGANIZACIÓN PANAMERICANA DE SALUD, 2003, *Sistemas de Información Geográfica en Salud Pública (SIG-SP)* (URL: <http://ais.paho.org/sigepi/index.asp>, consulta em 28-07-2008).
- PAINHO, M., 2006, [Aula da disciplina de Ciência e Sistemas de informação Geográfica do Mestrado em Ciência e Sistemas de informação Geográfica]. (Lisboa: ISEGI - UNL).
- RAKOTOMANANA, F., RANDREMANANA, R. V., RABARIJAONA, L. P., DUCHEMIN, J. B., RATOVO NJATO, J., ARIEY, F., RUDANT, J. P., e JEANNE, J., 2007, Determining areas that require indoor insecticide spraying using Multi Criteria Evaluation, a decision-support tool for malaria vector control programmes in the Central Highlands of Madagascar. *International Journal of Health Geographics*, 6:2 (URL: <http://www.ij-healthgeographics.com/content/6/1/2>, consulta em 1-3-08).
- RAMAKRISHNAN, R., e GEHRKE, J., 2000, *Database Management Systems* (2^a ed.) (McGraw-Hill).
- REGS, 1999, *EpiAnalyst for ArcView® GIS* (URL: <http://www.phrl.org/REGS/Info%20EpiAnalyst.htm>, consulta em 23-6-2008).

- REINHARDT, M., ELIAS, J., ALBERT, J., FROSCHE, M., HARMSSEN, D., e VOGEL, U., 2008, EpiScanGIS: an online geographic surveillance system for meningococcal disease. *International Journal of Health Geographics*, 7:33 (URL: <http://www.ij-healthgeographics.com/content/7/1/33>, consulta em 8-7-2008).
- RIORDAN, R., 2005, *Designing Effective Database Systems* (New Jersey: Addison Wesley Professional).
- ROGERSON, P. A., 1997, Surveillance system for monitoring the development of spatial patterns. *Statistics in Medicine*, 16, pp. 2081-2093 (Sussex: John Wiley & Sons) (URL: <http://www.ncbi.nlm.nih.gov/pubmed/9308133>, consulta em 11-7-2008).
- SANTANA, P., 2005, *Geografias da saúde e do desenvolvimento, evolução e tendências em Portugal* (Coimbra: Edições Almedina).
- SATSCAN, 2005, *SaTScan. Software for the spatial, temporal, and space-time scan statistics* (URL: <http://www.satscan.org/>, consulta em 2-6-2008).
- SCOTCH, M., BAMBANG, P., GADD, C.S. e SHARMA, R.K., 2006, Exploring the role of GIS during community health assessment problem solving: experiences of public health professionals. *International Journal of Health Geographics*, 5:39 (URL: <http://www.ij-healthgeographics.com/content/5/1/139>, consulta em 10-10-2006).
- SHUAI, J., BUCK, P., SOCKETT, P., ARAMINI, J., e POLLARI, F., 2006, A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada. *International Journal of Health Geographics*, 5:17 (URL: <http://www.ij-healthgeographics.com/content/5/1/17>, consulta em 5-7-2008).
- SILMAN, A. J., e MCFARLANE, G. J., 2004, *Epidemiological Studies: a practical guide* (2ª ed.) (Cambridge: Cambridge University Press).

- STEHMAN, S.V., 1999, Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, **20**, pp. 2423-2441 (URL: <http://www.informaworld.com/smpp/content~content=a713859829~db=all>, consulta em 16-4-2006).
- TAKAHASHI, K., KULLDORFF, M., TANGO, T., e YIH, K., 2008, A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics*, **7**:14 (URL: <http://www.ij-healthgeographics.com/content/7/1/14>, consulta em 5-7-2008).
- TRAN, A., PONÇON, N., TOTY, C., LINARD, C., GUIB, H., FERRÉ, J., SEEN, D., ROGER, F., ROCQUE, S., FONTENILLE, D., e BALDET, T., 2008, Using remote sensing to map larval and adult populations of *Anopheles hyrcanus* (Diptera: Culicidae) a potential malaria vector in Southern France. *International Journal of Health Geographics*, **7**:9 (URL: <http://www.ijhealthgeographics.com/content/7/1/9>, consulta em 01-03-2008).
- YI, Q., HOSKINS, R. E., HILLRINGHOUSE, E. A., SORENSEN, S. S., OBERLE, M. W., FULLER, S. S., e WALLACE, J. C., 2008, Integrating open-source technologies to build low-cost information systems for improved access to public health data. *International Journal of Health Geographics*, **7**:29 (URL: <http://www.ij-healthgeographics.com/content/7/1/29>, consulta em 5-7-2008).
- WALLER, L. A., e GOTWAY, C. A., 2004, *Applied Spatial Statistics for Public Health Data* (New Jersey: John Wiley & Sons).
- WARD, M. H., NUCKOLS, J. R., WEIGEL, S. J., MAXWELL, S. K., CANTOR, K. P., MILLER, R. S., 2000, Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System. *Environmental Health Perspectives*, **108**, pp. 5-12. (URL: <http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1637858&blobtype=pdf>, consulta em 5-7-2008).

WORLD HEALTH ORGANIZATION, 2005, *Avian Influenza: Assessing the Pandemic Threat* (URL: <http://www.who.int/csr/disease/influenza/H5N1-9reduit.pdf>, consulta em 3-4-2007).

ANEXOS

Anexo 1 - Especificações técnicas das tabelas da base de dados do SSES

EURO_URGENCIAS_H (SEQ_URGENCIA, COD_INSTITUICAO, DATA_REGISTO, DATA_NASCIMENTO, SEXO, COD_RESIDENCIA, COD_NATURALIDADE, COD_MOTIVO, COD_DESTINO). Afluência a urgências hospitalares – cada registo armazena uma consulta de urgência hospitalar.

SEQ_URGENCIA – chave primária, número sequencial único que identifica a consulta

COD_INSTITUICAO – Chave estrangeira, hospital onde a pessoa foi atendida

DATA_REGISTO – Data da ida à urgência

DATA_NASCIMENTO – Data de nascimento da pessoa atendida

SEXO - Chave estrangeira, sexo da pessoa atendida

COD_RESIDENCIA – Chave estrangeira, freguesia de residência da pessoa atendida

COD_NATURALIDADE – Chave estrangeira, país de origem da pessoa atendida

COD_MOTIVO – Chave estrangeira, motivo de ida à consulta

COD_DESTINO – Chave estrangeira, encaminhamento dado à pessoa após a consulta

EURO_URGENCIAS_H1 (SEQ_URGENCIA, COD_INSTITUICAO, DATA_REGISTO, DATA_NASCIMENTO, SEXO, COD_RESIDENCIA, COD_NATURALIDADE, COD_MOTIVO, COD_DESTINO, IDADE). Tabela idêntica à anterior, com a diferença de incluir um campo suplementar, denominado IDADE, que contém a idade da pessoa atendida, a qual é calculada no âmbito de processos que correm diariamente na base de dados.

EURO_URGENCIAS_C (Estrutura idêntica à tabela EURO_URGENCIAS_H).

Afluência a urgências de centros de saúde – cada registo armazena uma consulta de urgência em centro de saúde.

EURO_URGENCIAS_C1 (Estrutura idêntica à da tabela EURO_URGENCIAS_H1)

Afluência a urgências de centros de saúde – cada registo armazena uma consulta de urgência em centro de saúde incluindo um campo com cálculo da idade da pessoa atendida.

EURO_INTERNAMENTOS (SEQ_INTERNAMENTO, COD_INSTITUICAO, DATA_ALTA, DATA_INTERNAMENTO, DATA_NASCIMENTO, SEXO, COD_DESTINO). Discriminação

de episódios de internamentos hospitalares, correspondendo cada registo a um episódio de internamento hospitalar.

SEQ_INTERNAMENTO – Chave primaria, número sequencial único que identifica o internamento

COD_INSTITUICAO – Chave estrangeira, hospital onde a pessoa foi internada

DATA_ALTA – Data em que termina o período de internamento

DATA_INTERNAMENTO - Data de início do período de internamento

DATA_NASCIMENTO – Data de nascimento da pessoa internada

SEXO - Chave estrangeira, sexo da pessoa internada

COD_DESTINO – Chave estrangeira, encaminhamento dado à pessoa após internamento

EURO_INSTITUICOES (*CODIGO*, *COD_HOSP*, *COD_CS*, *COD_EXT*, *DESIG*, *MODO*, *ACTIVO*, *COD_REG*, *COD_SRS*). Características dos hospitais e centros de saúde incluídos no sistema

CODIGO - chave primária, código único que identifica o centro de saúde

COD_HOSP - subcódigo do hospital

COD_CS – subcódigo do centro de saúde

COD_EXT – subcódigo que identifica uma extensão do centro de saúde

DESIG – Nome do centro de saúde ou extensão

MODO – código do tipo (centro de saúde ou extensão)

ACTIVO – estado operacional do centro de saúde (aberto ou fechado)

COD_REG - chave estrangeira, região de saúde (NUTII) em que se localiza

COD_SRS - chave estrangeira, sub-região de saúde (Distrito) em que se localiza

EURO_RESIDENCIAS (*COD_RESIDENCIA*, *COD_DIST*, *COD_CONC*, *COD_FREG*, *COD_PAIS*, *MODO*, *DESC_RESIDENCIA*). Dados que definem o local de residência e naturalidade das pessoas atendidas.

COD_RESIDENCIA - Chave primária, local de residência da pessoa atendida

COD_DIST - subcódigo que identifica distritos de Portugal

COD_CONC - subcódigo que identifica concelhos de Portugal

COD_FREG, subcódigo que identifica freguesias de Portugal

COD_PAIS subcódigo que identifica países

MODO – Letra que identifica o tipo de unidade territorial interrogada (e.g. P – País)

DESC_RESIDENCIA – Nome da unidade territorial

EURO_DESTINOS_C (COD_DESTINO, DESC_DESTINO).

Encaminhamento após ida à urgência de centro de saúde.

COD_DESTINO – Chave primária, encaminhamento dado à pessoa após ida urgência

DESC_DESTINO – Descrição do tipo de encaminhamento

EURO_DESTINOS_H (COD_DESTINO, DESC_DESTINO).

Encaminhamento após ida a urgência hospitalar ou internamentos; estrutura idêntica à tabela anterior.

EURO_MOTIVOS_C (COD_MOTIVO, DESC_MOTIVO).

Motivo das idas às urgências de centros de saúde

COD_MOTIVO – Chave primária, motivo da ida urgência

DESC_MOTIVO – Descritivo do motivo

EURO_MOTIVOS_H (COD_MOTIVO, DESC_MOTIVO).

Discrimina o motivo das idas às urgências de hospitais, estrutura idêntica à tabela anterior

EURO_SEXO (COD_SEXO, DESC_SEXO).

Sexo da pessoa que ocorre às urgências

COD_SEXO – Chave primária, sexo da pessoa

DESC_SEXO – Descritivo do sexo

Anexo 2 – Excerto da sequência de execução das tarefas, Job 1

Explicação da sequência de execução das tarefas, incluindo os blocos de código (*Programs*) PL/SQL e as tabelas que sofrem operações. É apenas explicada uma parte da sequência, sendo a forma de cálculo das restantes partes semelhante em tudo, excepto no que respeita aos valores finais, que se referem a classes etárias. É também apresentada no final uma imagem ilustrativa da modelação do *Job 1* em *Mindjet Mind Manager*, sendo assinalada a vermelho, a parte em foco neste texto.

SIVEADMIN_EURO_URGENCIAS_H

Tabela de base, actualizada diariamente a partir do servidor de recolha; contém os registos, por pessoa atendida, de procura diária em urgências hospitalares

FLUXO1_P1

Apaga previamente o conteúdo da tabela SIVEADMIN.EURO_URGENCIAS_H1

```
begin
DELETE FROM SIVEADMIN.EURO_URGENCIAS_H1;
commit;
end;
```

FLUXO1_P2

Transfere o conteúdo da tabela SIVEADMIN.EURO_URGENCIAS_H para a tabela SIVEADMIN.EURO_URGENCIAS_H1 e calcula um campo *idade* suplementar

```
begin
INSERT INTO SIVEADMIN.EURO_URGENCIAS_H1
SELECT
SEQ_URGENCIA,
COD_INSTITUICAO,
DATA_REGISTO,
DATA_NASCIMENTO,
```

```

SEXO,
COD_RESIDENCIA,
COD_NATURALIDADE,
COD_MOTIVO,
COD_DESTINO,
(DATA_REGISTO-DATA_NASCIMENTO)/365 IDADE
FROM SIVEADMIN.EURO_URGENCIAS_H
WHERE DATA_REGISTO >= '2008-01-01'
AND (DATA_REGISTO-DATA_NASCIMENTO)/365 >= 0
AND (DATA_REGISTO-DATA_NASCIMENTO)/365 <= 110;
commit;
end;

```

SIVEADMIN_EURO_URGENCIAS_H1

Tabela transitória; estrutura semelhante à tabela

SIVEADMIN_EURO_URGENCIAS_H, acrescida do campo *idade*

FLUXO1_P3

Apaga o conteúdo da tabela SIVEADMIN.EURO_H_DIST_TI_V2;

```

begin
DELETE FROM SIVEADMIN.EURO_H_DIST_TI_V2;
commit;
end;

```

FLUXO1_P4

Calcula o conteúdo da tabela SIVEADMIN.EURO_H_DIST_TI_V2

```

begin
INSERT INTO SIVEADMIN.EURO_H_DIST_TI_V2
    SELECT DATA_REGISTO,
    COUNT (SEQ_URGENCIA),

```

```

COUNT (DECODE(COD_SRS,1,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,2,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,3,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,4,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,5,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,6,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,7,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,8,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,9,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,10,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,11,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,12,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,13,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,14,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,15,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,16,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,17,SEQ_URGENCIA)),
COUNT (DECODE(COD_SRS,18,SEQ_URGENCIA))
FROM SIVEADMIN.EURO_INSTITUICOES,
     SIVEADMIN.EURO_URGENCIAS_H1
WHERE  CODIGO=COD_INSTITUICAO
AND    COD_MOTIVO = 5
GROUP BY DATA_REGISTO
ORDER BY DATA_REGISTO;
commit;
end;

```

SIVEADMIN.EURO_H_DIST_TI_V2

Tabela transitória; contém totais de procura diária de serviços de urgência hospitalares por distrito, todas as idades, todos os sexos

FLUXO3_P1

Apaga o conteúdo da tabela SIVEADMIN.TOTDIST_UDIAH;

```
begin
DELETE FROM SIVEADMIN.TOTDIST_UDIAH;
commit;
end;
```

FLUXO3_P2

Calcula o conteúdo da tabela SIVEADMIN.TOTDIST_UDIAH

```
begin
INSERT INTO SIVEADMIN.TOTDIST_UDIAH (TOT_UDIA,COD_SRS)
SELECT COUNT(T1.SEQ_URGENCIA) AS TOT_UDIA, T2.COD_SRS
FROM SIVEADMIN.EURO_URGENCIAS_H1 T1,SIVEADMIN.EURO_INSTITUICOES T2
WHERE T1.COD_INSTITUICAO=T2.CODIGO
AND DATA_REGISTO=(SELECT MAX(DATA_REGISTO) FROM
SIVEADMIN.EURO_URGENCIAS_H1)
AND COD_MOTIVO = 5
GROUP BY T2.COD_SRS;
commit;
end;
```

SIVEADMIN.TOTDIST_UDIAH

Tabela transitória para alimentar mapas coropletos da procura diária de serviços de urgência hospitalares por distrito, todas as idades, todos os sexos. Totais referentes ao dia anterior.

FLUXO3_P2a

Calcula o conteúdo da tabela SIVEADMIN.TOTDIST_UDIAHV2

```
begin
MERGE INTO SIVEADMIN.TOTDIST_UDIAHV2 a
  USING SIVEADMIN.TOTDIST_UDIAH b
    ON (a.COD_SRS = b.COD_SRS)
  WHEN MATCHED THEN
```

```

UPDATE SET a.TOT_UDIA = b.TOT_UDIA;
commit;
end;

```

SIVEADMIN.TOTDIST_UDIAHV2

Tabela transitória similar à anterior, destinada a resolver um problema: uma vez que a tabela anterior é povoada após esmagamento (todo o conteúdo é apagado), por vezes, se houver interrupção no envio de dados para um distrito, o mapa final não o apresentará, ficando no seu lugar um espaço vazio. Para tal, esta tabela foi gerada previamente, sendo *actualizada* (comando MERGE) ao invés de *esmagada*.

FLUXO3_P3

Apaga o conteúdo da tabela SIVEADMIN.TOTDIST_UDIA2H;

```

begin
DELETE FROM SIVEADMIN.TOTDIST_UDIA2H;
commit;
end;

```

FLUXO3_P4

Calcula o conteúdo da tabela SIVEADMIN.TOTDIST_UDIA2H

```

begin
INSERT INTO SIVEADMIN.TOTDIST_UDIA2H
SELECT T1.COD_SRS,(T1.TOT_UDIA/T2.HM)*100000 AS TOT_UDIAP
FROM SIVEADMIN.TOTDIST_UDIAHV2 T1,SIVEADMIN.POPF2005 T2
WHERE T1.COD_SRS=T2.COD_SRS;
commit;
end;

```


SIVEADMIN.TOTDIST_UDIA2H

Tabela transitória para alimentar mapas coropletos da procura diária de serviços de urgência hospitalares por distrito, todas as idades, todos os sexos, ponderados pela população residente e por 100 000 habitantes.

FLUXO3_P4V2

Calcula o conteúdo da tabela SIVEADMIN.TOTDIST_UDIA3H

```
begin
MERGE INTO SIVEADMIN.TOTDIST_UDIA3H a
USING SIVEADMIN.TOTDIST_UDIA2H b
ON (a.COD_SRS = b.COD_SRS)
WHEN MATCHED THEN
UPDATE SET a.TOT_UDIAP = b.TOT_UDIAP;
COMMIT;
end;
```

SIVEADMIN.TOTDIST_UDIA3H

Tabela final para alimentar mapas coropletos da procura diária de serviços de urgência hospitalares por distrito, todas as idades, todos os sexos, ponderados pela população residente e por 100 000 habitantes. É igualmente povoada por actualização (comando MERGE), constituindo um artifício para resolver um problema com a produção automática da escala de cores do mapa final; inclui um 19º distrito (fictício), cujo valor se encontra sempre a zero, provocando o resultado desejado no mapa final.

SIVEADMIN.POPF2005

Tabela contendo a população residente por distrito.

SIVEADMIN_EURO_INSTITUICOES

Tabela contendo as características das instituições (hospitais e centros de saúde e respectivo código de unidade territorial)

